2012

Characterization of the Promoter and Insulator Activities of the Scs' Insulator

Yu Ge
Louisiana State University and Agricultural and Mechanical College, yge2@lsu.edu

CHARACTERIZATION OF THE PROMOTER AND INSULATOR ACTIVITIES OF THE SCS' INSULATOR

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements of the degree of
Master of Science

in

The Department of Biological Science

by
Yu Ge
B.S., Sichuan Agricultural University, 2009
August 2012

ACKNOWLEDGMENTS

I would like to thank Dr. Craig Hart, my major professor, for his patience and guidance and also Dr. Patrick DiMario, Dr. David Donze, and Dr. Joomyeong Kim, my dissertation committee members, for their scientific advice and time.

I would also like to thank members in Dr. Craig Hart lab and Dr. Patrick DiMario lab: Avva Satya, Shraddha Shrestha, Brandon Finnorn, Trace Favre, Allison James and Yubo Wang for their great help in my research and life.

I would also like to acknowledge the Department of Biological Science and Graduate School for their financial support for my graduate study.

Last but not the least, I would like to thank my parents and my girlfriend for their love and encourage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AttB	Bacterial Attachment Site
AttP	Phage Attachment Site
B site	Low BEAF Affinity Binding Site
BEAF	Boundary Element Associated Factor
BED fingers	C2H2 Zinc Finger DNA Binding Domains
BX-C	Bithorax Complex
CP190	Centrosomal Protein 190 kDa
CTCF	CCCTC Binding Factor
dTopors	*Drosophila* Topoisomerase I-interacting RS Proteins
DMR	DNA Methylated Region
D site	High BEAF Affinity Binding Site
DTT	Dithiothreitol
EMSA	Electrophoretic Mobility Shift Assay
GAF	GAGA Factor
HAT	Acetyltransferase
HMT	Histone Methyltransferase

H3C	High-resolution Chromatin Conformation Capture
ICR	Imprinted Control Region
Mod (mdg4)	*Modifier of mdg4*
PRE	Polycomb Response Element
SB	Squishing Buffer
SCS	Special Chromatin Structure
SCS’	Special Chromatin Structure’
Su(Hw)	Suppressor of Hairy-wing
USF1	Upstream Stimulatory Factor 1
USF2	Upstream Stimulatory Factor 2
VEZF1	Vascular Endothelial Zinc Finger 1
ZW5	Zeste-white 5

ABSTRACT

Eukaryotic gene expression is regulated by regulatory elements. Insulators are one class of regulatory sequences. They can prevent heterochromatin from spreading into euchromatin, block distal enhancer activity, or both. Insulators function by recruiting several associated proteins. The scs' insulator, which is located at one end of the 87A7 *hsp* 70 locus, is bound by BEAF. In the *Drosophila* genome, more than 85 % of BEAF binding sites are found within 300 bp of transcription start sites, suggesting BEAF may be involved in promoter function. Based on previous insulator position effect assays, two sequences were found to be necessary for full scs' insulator function. One is the high affinity BEAF binding site called "D", and the other is a 20 bp sequence called "LS4". In order to identify the minimal scs' insulator sequence, several scs' derivative sequences containing the LS4 region and D site will be tested using a site specific integration system. In this system, ΦC31 integrase mediates specific *attB* site integration into transgenic *attP* sites in the *Drosophila* genome. Two *attP* fly lines with strong position effects are needed to test scs' derivative sequences. Currently, 4 out of 12 transgenic *attP* fly lines show slight position effects. Making new *attP* landing site flies is in progress to find at least two lines that have strong position effects. Promoter activity assays in S2 cells demonstrated that the scs' M fragment possesses promoter activity in addition to insulator activity. Electrophoretic mobility shift assays (EMSAs) detected proteins that bind to the LS4 region, and BEAF binding facilitates this binding. DNA affinity chromatography was performed to purify LS4 binding proteins. After two rounds of purification, several candidate protein bands were identified. Further characterization needs to be done to confirm these proteins bind to the LS4 region. If confirmed, the proteins will be identified by mass spectrometry.

CHAPTER 1
INTRODUCTION

1.1 INSULATORS AND INSULATOR PROTEINS

Gene expression of eukaryotic organisms is ensured by considerable regulatory elements, which can be located at remarkable distances from their promoters (Bulger and Groudine, 2010; Tchurikov et al., 2009). Insulators or boundary elements are DNA elements which regulate the interaction of regulatory elements and further affect chromatin architecture and organization. Based on previous studies, insulators are classically featured with two experimental properties, barrier activity and enhancer-blocking activity (Gaszner and Felsenfeld, 2006). Barrier activity involves the protection against heterochromatin-mediated silencing; enhancer-blocking activity refers to the capacity to prevent communication between regulatory elements and promoters. In some insulators, such as *gypsy* and *Idefix*, these two activities can not be separated (Brasset et al., 2010; Roseman et al., 1993); others have either enhancer-blocking or barrier activity. For example, there are five DNase I hypersensitive sites within the chicken β globin 5'HS4 insulator. CTCF binds to one of these DNase I hypersensitive sites, and is essential and necessary for enhancer blocking activity, but has no effect on barrier activity. The other four sites are necessary for barrier activity and dispensable for enhancer blocking activity. It indicates barrier activity and enhancer blocking activity are separate in some cases (Recillas-Targa et al., 2002; West et al., 2004).

Insulators and their binding proteins are well characterized in *Drosophila* and mammals. ChIP-chip data of the *Drosophila* genome demonstrated there are more than 14000 insulator protein binding sites in the genome (Negre et al., 2010). In *Drosophila*, insulators like *gypsy*, *Fab 7*, *Fab 8*, *scs'* and *scs* are well studied and bound by Suppressor of hairy-wing [Su(Hw)], GAGA factor (GAF), *Drosophila* CTCF (dCTCF), Boundary element associated factor (BEAF)

and Zeste-white 5 (ZW5), respectively. Moreover, centrosomal protein 190 (CP190), Mod (mdg4)2.2, *Drosophila* topoisomerase I- interacting RS proteins (dTopors) are reported to facilitate insulator function (Bushey et al., 2009; Capelson and Corces, 2005; Gerasimova et al., 1995; Pai et al., 2004).

The *gypsy* insulator, one of the well characterized insulators, is a 340 bp sequence from *gypsy* retrotransposon in *Drosophila melanganster*, which was identified through the study of mutant phenotype in *Drosophila* (Modolell et al., 1983). It was found that mutation of the second site modifier gene, later identified as coding Su(Hw), rescued the mutant phenotype caused by the insertion of a *gypsy* retrotransposon (Rutledge et al., 1988), suggesting Su(Hw) genetically interacts with *gypsy* sequence. Su(Hw) has twelve zinc fingers and binds to a 12-bp motif YRYTGCATAYYY (Y-Pyrimidine, R-Purine) in *gypsy* insulator (Parnell et al., 2006). Genomic mapping of Su(Hw) binding sites in *Drosophila* revealed that the binding sites could extend to 20-bp motif which allows variation in TGCATA core region (Adryan et al., 2007). The *gypsy* retrotransposon contains twelve tightly clustered Su(Hw) binding sites separated by A/T rich sequences (Dorsett, 1993; Spana and Corces, 1990; Spana et al., 1988). The clustered Su(Hw) binding sites alone cause the same mutagenic phenomenon as that of *gypsy* retrotransposon and block enhancer activity in enhancer- blocking assays, further proving that Su(Hw) plays a vital role in mutagenesis by insertion of *gypsy* and that this region functions as an insulator (Geyer and Corces, 1992; Holdridge and Dorsett, 1991; Scott et al., 1999). Additionally, mutations of these Su(Hw) binding sites dramatically compromise the mutagenesis caused by *gypsy* retrotransposon (Peifer and Bender, 1988; Smith and Corces, 1992). Other studies have demonstrated that the inserted gypsy sequence also functions as a barrier to prevent chromosome position effects in the *Drosophila* genome (Markstein et al., 2008). Besides

reversing gypsy-induced phenomenon, most Su(Hw) mutants have abnormal oogenesis in which oocytes are lost by apoptosis during mid-oogenesis, suggesting Su(Hw) may be involved in female germline development (Baxley et al., 2011; Soshnev et al., 2012). Interestingly, insulator activity is lost when two copies of the *gypsy* insulator, rather than a single one, are inserted between a promoter and an enhancer, which is called insulator bypass (Muravyova et al., 2001). In addition to insulator activity in *Drosophila*, interestingly, it is reported that transgenes flanked by *gypsy* insulators, along with coexpression of Su(Hw), are not subject to chromosomal position effects no matter where the insertion site is in the genome of *Arabidopsis thaliana* (She et al., 2010). It provides a novel method to obtain precise transgene expression in plants. Recently, a body of evidence experimentally demonstrated that CP190, Mod (mdg4)2.2 and dTopors all interact with Su(Hw) at the *gypsy* insulator to facilitate insulator function in *Drosophila* (Bushey et al., 2009; Negre et al., 2010; Pai et al., 2004; Petersen et al., 1994; Ramos et al., 2011).

Fab-7, along with *Mcp* and *Fab-8*, are cis-regulatory sequences that regulate expression of the *Abd- B* gene in the *Drosophila* bithorax complex (BX-C) (Barges et al., 2000; Gyurkovics et al., 1990; Karch et al., 1994). *Fab-7* is one of the best characterized regulatory elements in the BX-C region. It behaves as an insulator to separate the *iab-6* and *iab-7* regulatory regions which control abdominal segment development (Mihaly et al., 1997). It is reported that *Fab-7* mutations cause the fusion of the *iab-6* and *iab-7* regulatory regions, resulting in the transformation of abdominal segment 6(A6) into A7 by affecting *Abd-B* expression. There are nine consensus GAF binding sites- (GAGAG)- within *Fab-7*. Enhancer-blocking assays demonstrated that mutations of GAF binding sites in *Fab-7* suppress its boundary activity, with some sites being more important (Schweinsberg et al., 2004). GAF, encoded by *Trithorax-like* gene (Farkas et al., 1994), was originally observed as a transcription factor (Biggin and Tjian,

1988). There are two isoforms of GAF: GAF519aa and GAF581aa in *Drosophila* (Soeller et al., 1993). They share the same N-terminus, including a DNA binding domain and BTB/POZ domain; they differ in the C-terminus which is rich in glutamine. It is believed that the BTB/POZ domain is responsible for homologous and heterologous protein: protein interactions, whereas the C-terminus contributes to homologous protein: protein interactions (Faucheux et al., 2003; Mishra et al., 2003). GAF mediates boundary function not only in the *Fab-7* boundary domain, but also within the Eve promoter and SF1 boundary elements (Belozerov et al., 2003; Ohtsuki and Levine, 1998).

Fab-8, located downstream of adjacent *Fab-7*, prevents initiation elements and a Polycomb Response Element (PRE) in the *iab-8* domain from affecting the *iab-7* regulatory region (Barges et al., 2000). Unlike *Fab-7*, *Fab-8* is bound by the insulator protein *Drosophila* CTCF (dCTCF), which is homologous to the only insulator protein discovered in vertebrates (Moon et al., 2005). dCTCF possesses twelve zinc fingers. The *dCTCF* gene is essential, and mutations cause abdominal hometic phenotypes (Gerasimova et al., 2007). In vertebrates, CTCF contains a central eleven zinc finger DNA binding domain. The insulator activity of CTCF was originally found at the 5' HS4 and 3' HS1 insulators which developmentally regulate mouse β-globin expression (Farrell et al., 2002). Enhancer- blocking assays showed that CTCF is essential for the insulator activity (Bell et al., 1999). Besides insulator activity, CTCF plays a vital role in diverse cellular processes. It is embryonic lethal in CTCF knock-out mice (Heath et al., 2008).

The 87A7 *hsp 70* locus is flanked by two insulators that form special chromatin structures. One is called scs, the other is called scs' (Udvardy et al., 1985). The scs insulator is able to protect transgenes from chromosomal position effects and block enhancer activity in enhancer blocking assays (Kellum and Schedl, 1992). ZW5, a zinc finger protein, binds to a 24-bp

sequence in scs which is required for insulator activity. Genetic studies showed that deletion of ZW5 was recessive lethal which makes genetic analysis difficult (Gaszner et al., 1999; Maeda and Karch, 2007). Boundary element associated factors (BEAFs), which our lab is interested in, bind to scs'. More details of scs' and BEAF are given in the next section.

In addition to insulator proteins mentioned above, some proteins are reported to be recruited by the above insulator binding proteins. Mod(mdg4), a BTB/POZ protein capable of oligomerization and encoded by the *Modifier of mdg4* [*Mod(mdg4)*] gene, has at least 27 different isoforms. It was reported that only Mod(mdg4)-67.2, also called Mod(mdg4)2.2, was involved in insulator activity in enhancer-blocking assays (Gerasimova et al., 1995). More evidence proved that Mod (mdg4)-67.2 played a role in the establishment of early embryonic epigenetic marks (Gerasimova and Corces, 1998). Moreover, Mod (mdg4)-67.2 interacts with Su(Hw) *in vivo* through its unique carboxy-terminal domain (Buchner et al., 2000; Gause et al., 2001). CP190 contains C2H2 zinc finger domains, BTB/POZ domains and a glutamine-rich C-terminus. It was originally found to be associated with centrosomes by binding to β-tubulin in mitosis and with chromosomes in interphase, respectively (Jimenez and Goday, 1993; Raff et al., 1993). However, studies showed that CP190 is essential for fly viability but is dispensable for normal centrosomal function (Butcher et al., 2004). CP190 was colocalized with the insulator proteins dCTCF, Su(Hw) and BEAF in *Drosophila* Kc167 cells and Mbn2 cells (Bushey et al., 2009; Negre et al., 2010). CP190 does not bind to DNA directly. At *gypsy* elements, it interacts with Su(Hw) and Mod(mdg4)-67.2. Evidence showed that CP190 levels are low after heat shock, and this correlates with lower levels of global gene expression in *Drosophila* (Wood et al., 2011). In addition to involvement in insulator activity, myosin organization was disrupted in *CP190*

mutant embryos, suggesting that CP190 is involved in the regulation of myosin function in *Drosophila* (Chodagam et al., 2005).

1.2 SCS' AND BOUNDARY ELEMENT ASSOCIATED FACTORS (BEAF)

The special chromatin structures scs and scs' flank the *Drosophila* 87A7 *hsp 70* locus (Udvardy et al., 1985). These elements with two sets of DNaseI hypersensitive sites within each sequence are the first two boundary elements defined by enhancer-blocking assays (Kellum and Schedl, 1991, 1992). Electrophoretic mobility shift assays (EMSAs) and footprinting assays led to the discovery of two BEAF binding sites within scs' sequences which share variably arranged CGATA motifs (Zhao et al., 1995). BEAF binding sites are indispensable for scs' insulator activity. In position independent expression assays, an scs' derivative M2 fragment, a dimer containing two copies of one BEAF binding site, functions as a boundary element as effectively as scs'; whereas point mutations in the BEAF binding sites eliminate the boundary activity (Cuvier et al., 1998). Later experiments showed that M2 fragment was unable to block position effects in three out of four tested transgenic fly lines when BEAF was mutated, suggesting BEAF is essential for insulator activity (Roy et al., 2007). BID, a dominant-negative form of BEAF, has a BEAF self-interaction domain but lacks a DNA binding domain. Expression of BID is lethal in early embryogenesis, indicating BEAF is required in early development. Expression of BID in eye imaginal discs and salivary glands resulted in a rough-eye phenotype and the disruption of polytene chromatin structure, respectively. Both phenotypes could be rescued by a third copy of the BEAF gene, suggesting BEAF is involved in maintaining chromatin structure and dynamics (Gilbert et al., 2006; Roy et al., 2007).

There are two isoforms of BEAF in *Drosophila melanogaster*: 32A and 32B. They share the same C-terminal domain which mediates BEAF-BEAF interactions, and differ in their N-

terminal domains which contain C2H2 zinc finger DNA binding domains (BED fingers) (Aravind, 2000; Hart et al., 1997). BEAF knock-out experiments demonstrated that 32B alone is sufficient for fly survival (Roy et al., 2007). It was reported that DREF, characterized as a transcription factor, competed with BEAF for binding to some DNA sequences (Hart et al., 1999). Immunostaining of polytene chromosomes showed BEAF was distributed throughout the whole genome (Zhao et al., 1995). Recent studies have found that more than 85% of over 1800 BEAF binding sites are located within 300bp from transcription start sites (TSSs). BEAF knockout affects associated gene expression, suggesting BEAF may play roles both in insulator activity and transcription (Bushey et al., 2009; Jiang et al., 2009).

1.3 MECHANISMS OF INSULATORS

Although insulators or boundary elements were known for decades, molecular mechanisms are still not well understood. However, various models have been proposed to explain enhancer-blocking and barrier activities. These models are not mutually exclusive.

In barrier activity, insulators possibly prevent heterochromatin from spreading into euchromatin region by recruiting gene-activating or histone-modifying factors. Characteristics of heterochromatin are low levels of histone acetylation and high levels of H3 Lys9 (H3K9) and H3 Lys27 (H3K27) methylation, and enriched for heterochromatin protein 1 (HP1). Euchromatin has a high level of histone acetylation and H3 Lys4 (H3K4) and H3 Lys 79 (H3K79) methylation. The chicken β globin 5'HS4 insulator prevents heterochromatin from spreading by upstream stimulatory factor 1 (USF1) and upstream stimulatory factor 2 (USF2) binding to hypersensitive sites. The binding of USF1 and USF2 recruits a histone acetyltransferase (HAT) and a histone methyltransferase (HMT) which make this region highly acetylated and H3K4 methylated, thus preventing heterochromatin from spreading (Gaszner and Felsenfeld, 2006). Additionally,

transgene promoters flanked by chicken 5'HS4 insulators are protected from *de novo* DNA methylation by a novel chromatin barrier protein- vascular endothelial zinc finger 1 (VEZF1) (Dickson et al., 2010; Mutskov et al., 2002).

As to enhancer-blocking activity, there are two models proposed currently: promoter-decoy model and loop model. The promoter-decoy model suggests that insulators compete with promoters for interacting with enhancers. Nipped-B and chip proteins are identified as enhancer binding factors. Mutation of *chip* or *Nipped* magnified the effects of *gypsy* transposon insertion in *cut* gene, suggesting that they genetically interact with the insulator or insulator proteins (Morcillo et al., 1997; Rollins et al., 1999). In human K562 cells, an interposed chicken 5'HS4 insulator blocked HS2 enhancer activity and caused pol II and TBP to accumulate at the insulator site. Chromatin conformation capture (3C) found that the HS2 enhancer has a higher frequency of localization with the 5' HS4 insulator than with a promoter (Zhu et al., 2007). However, the loop model focuses on chromatin loops formed by insulator interaction with another insulator or other structures in the nucleus. It was reported that Su(Hw) binds to the *gypsy* insulator, and recruits CP190 and Mod (mdg4)-67.2. Mod (mdg4)-67.2 interacts with dTopors, which interacts with nuclear lamin. An insulator chromatin loop is formed by the communication between these proteins (Capelson and Corces, 2005; Ong and Corces, 2009). High-resolution chromatin conformation capture (H3C) method was used to show that two *gypsy* insulators spatially formed a loop to allow communication between polycomb response element (PRE) and a distal *mini-white* promoter (Comet et al., 2011). Besides *gypsy* insulators, scs and scs' form a chromatin loop *in vivo* by interactions between Zw5 and BEAF at the opposite ends of the 87A7 *hsp70* locus (Blanton et al., 2003). Mammalian *Igf2* (insulin-like growth factor 2) and *H19* imprinted genes are regulated by an ICR (imprinted control region), DMR (DNA methylated region) and

enhancer elements downstream of the *H19* gene. There are four CTCF binding sites in the ICR, and DNA methylation of these sites can inhibit CTCF binding. On the maternal allele, CTCF mediates the loop formation between ICR and DMR1 to block the interaction between the downstream enhancer and *Igf2* gene. On the paternal allele, ICR Methylation eliminates CTCF binding, which allows the downstream enhancer to activate Igf2 gene (Ong and Corces, 2009).

Although barrier models and enhancer-blocking models differ, all of the models suggest that several proteins bind to insulator sequences directly or indirectly to facilitate insulator function. The aggregates of insulator proteins are called "insulator bodies" (Gaszner and Felsenfeld, 2006; Golovnin et al., 2008). The loss of insulator proteins dramatically compromises insulator activity.

1.4 GLOBAL DISTRIBUTION OF INSULATOR PROTEINS

Since insulators are characterized by binding proteins, it is possible to explore their global distribution in the genome. Chromatin immunoprecipitation (ChIP) and microarray analysis (ChIP-chip) data indicated that Su(Hw) and dCTCF colocalize with CP190 in Kc cells and Mbn2 cells, which is consistent with the fact that CP190 was found in both insulator bodies (Mohan et al., 2007; Pai et al., 2004). Interestingly, BEAF is also found to colocalize with CP190 in both cells, although there is no evidence BEAF aggregates into insulator bodies. Su(Hw), dCTCF and BEAF together account for more than 85% of CP190 binding sites in the Kc cell genome, suggesting CP190 plays a crucial role in insulator activity. Additionally, CP190 is associated with H3 depletion at TSSs, suggesting CP190 binds to active promoters (Bartkuhn et al., 2009). Only 20% of Su(Hw) sites and 47% of dCTCF sites are located within 1kb of genes; however, around 70% of CP190 sites and 84% of BEAF sites are within 1kb of genes (Bushey et al., 2009). This distribution is consistent with the data from our lab, which found that more than 85% of BEAF sites were within 300bp of transcription start sites. In addition, it was reported that

polymerase negative elongation factor (NELF) was present at 40% of BEAF associated genes (Jiang et al., 2009). All of these suggest that BEAF may be involved in promoter activity as well as insulator activity. It has been reported that some transcription factors assist in both insulator and promoter activity. In *Schizosaccharomyces pombe*, the transcription factor TFIIIC complex without RNA polymerase (Pol) III, binds to B-box sequences within inverted repeat (IR) elements. These IR elements act as barriers to protect euchromatic regions from being affected by adjacent silenced mating-type loci (Noma et al., 2006). Also, tRNA genes and even TFIIIC binding sites possess both barrier and insulator activity in *Saccharomyces cerevisiae* (Donze and Kamakaka, 2001; Simms et al., 2008). Taken together, CP190 is shared by other major insulators in insulator function in *Drosophila*, and most of BEAF and CP190 sites are distributed close to TSSs, indicating BEAF and CP190 may play a role in transcription activation.

1.5 *DROSOPHILA* AS A MODEL ORGANISM

It is impractical to do research on human beings directly, considering many factors, like ethics, cost, time, feasibility etc. However, there are several useful model organisms, such as bacteria, yeasts, *Caenorhabditis elegans* (*C. elegans*), *Drosophila melanogaster*, mice and primates. Compared with other model organisms, *Drosophila* has some advantages which make it an indispensable and powerful tool for research.

First of all, fruit flies are eukaryotic organisms and relatively close to human beings genetically compared to bacteria or yeasts. Flies and humans share many homologous proteins and pathways, which makes flies suitable as a model system.

Second, fruit flies have a relatively short life cycle compared with many other animals. The life cycle of a fruit fly is around 3 weeks at 20 ℃, 10 days at 25 ℃. It is easy and inexpensive to maintain. In the lab, we put flies in polypropylene vials and feed them with live yeast. You don’t

need to buy an incubator to keep flies. Leaving them at room temperature is fine. Moreover, fly vials containing flies do not take up a lot of space. Taking all of this into consideration, a small lab could easily set up a fly system without excessive expense.

Third, fruit flies are simple compared to mammals, making them good for genomics and proteomics research. Besides, a lot of reporter genes and balancer chromosomes are available for *Drosophila* genetics, and most biochemical analysis methods could apply to flies. These advantages make *Drosophila* an excellent tool for genetic and biochemical studies.

CHAPTER 2
IDENTIFICATION OF MINIMAL SCS' SEQUENCES BY USING THE ΦC 31 SYSTEM

2.1 INTRODUCTION

As described in Chapter 1, scs' is the first characterized insulator which has a BEAF binding motif. Previous studies have demonstrated that scs' is defined by nuclease- hypersensitive sites arranged around a central nuclease-resistant region (Udvardy et al., 1985). Interestingly, two palindromic segments (CGATAnTATCG and CGATAnnnTATCG) form the core of high and low affinity BEAF binding sites (D site and B site, respectively). They are adjacent to the previously mapped resistant region of scs' (Figure 2.1).

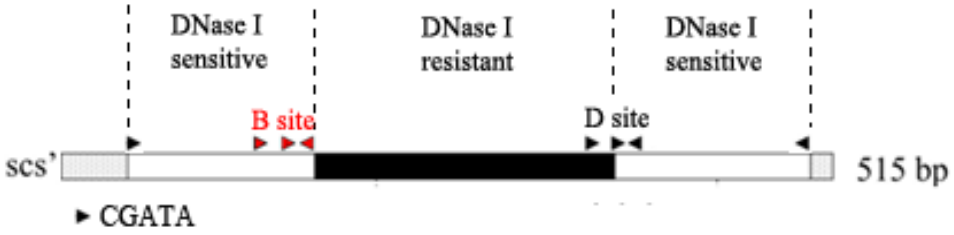

Figure 2.1 BEAF binding sites on scs'. The arrowheads represent the position and orientation of CGATA motifs. Two DNase I hypersensitive regions and two BEAF binding sites (B site and D site) are in scs'. The D site has a higher BEAF affinity than the B site.

A position effect assay was used to detect insulator activity. In this assay, the *mini-white* gene, in which no enhancer is included, is used as a reporter. The *mini-white* gene should be expressed at a basal level, resulting in low production of eye pigmentation and light eye color. In the assay, scs, which has a strong boundary function, is inserted downstream of *mini-white* to block the influence of downstream regulatory elements. Sequences of interest from scs' are placed upstream of *mini-white* to test for insulator activity (Figure 2.2, Figure 2.3). If inserted without an insulator in an active region in the genome, the *mini-white* gene is expressed at a high level, resulting in a dark eye color. In such a case, the insertion is subject to a strong position effect. However, if an upstream insulator is able to block the communication between upstream

positive regulatory elements and the *mini-white* gene, then the *mini-white* gene is expressed at a basal level, resulting in a light eye color. In other words, the insulator blocks the position effect (Figure 2.2).

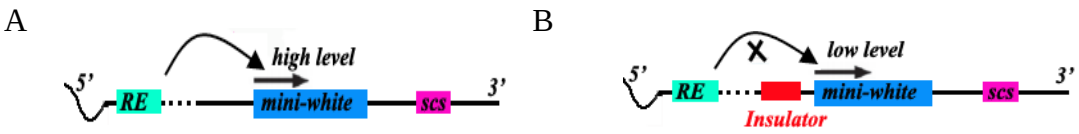

Figure 2.2 Diagram of position effect assays. RE is positive regulatory element, such as enhancer, in the *Drosophila* genome. The scs insulator is placed downstream of *mini-white* gene to block the influence from downstream regulatory elements.

A 225 bp sequence from scs' contains the D site, a high affinity BEAF binding site. This 225 bp sequence is named the M fragment (Figure 2.3). A dimer of the M fragment (M2), which has two D sites, functions as well as scs' in the position effect assay. 90% of fly lines with the M2 have light eyes, just as with scs'. Point mutations in the D site (M*2) eliminate BEAF binding and the insulator activity in the position effect assay. Only 30% of fly lines with M*2 show light eye colors, suggesting that the D site is essential for insulator function in the form of BEAF binding (Cuvier et al., 1998). The smaller dimer fragments S2 and X2, both derived from M2, reduced but did not eliminate insulator activity. This suggests that regions in M2 missing in S2 or X2 increase the insulator activity of the BEAF binding site (Figure 2.3, Table 2.1). This region was divided into six 20-bp mutated segments for a linker scanning analysis. Of these regions, named from LS1 to LS6, only the LS4 mutation weakened insulator function (Table 2.1). Taken together, the D site and LS4 region are essential for full scs' insulator function. It is currently unknown whether a monomer of the M fragment is sufficient for insulator activity and whether there are additional regions involved in insulator activity. To address this, we are interested in the following questions: 1). Does M containing a single D site works as well as M2 containing two D sites? 2). Are only the D site and LS4 region sufficient for full insulator activity? 3). Is

the spacing between the D site and the LS4 region important for insulator function?

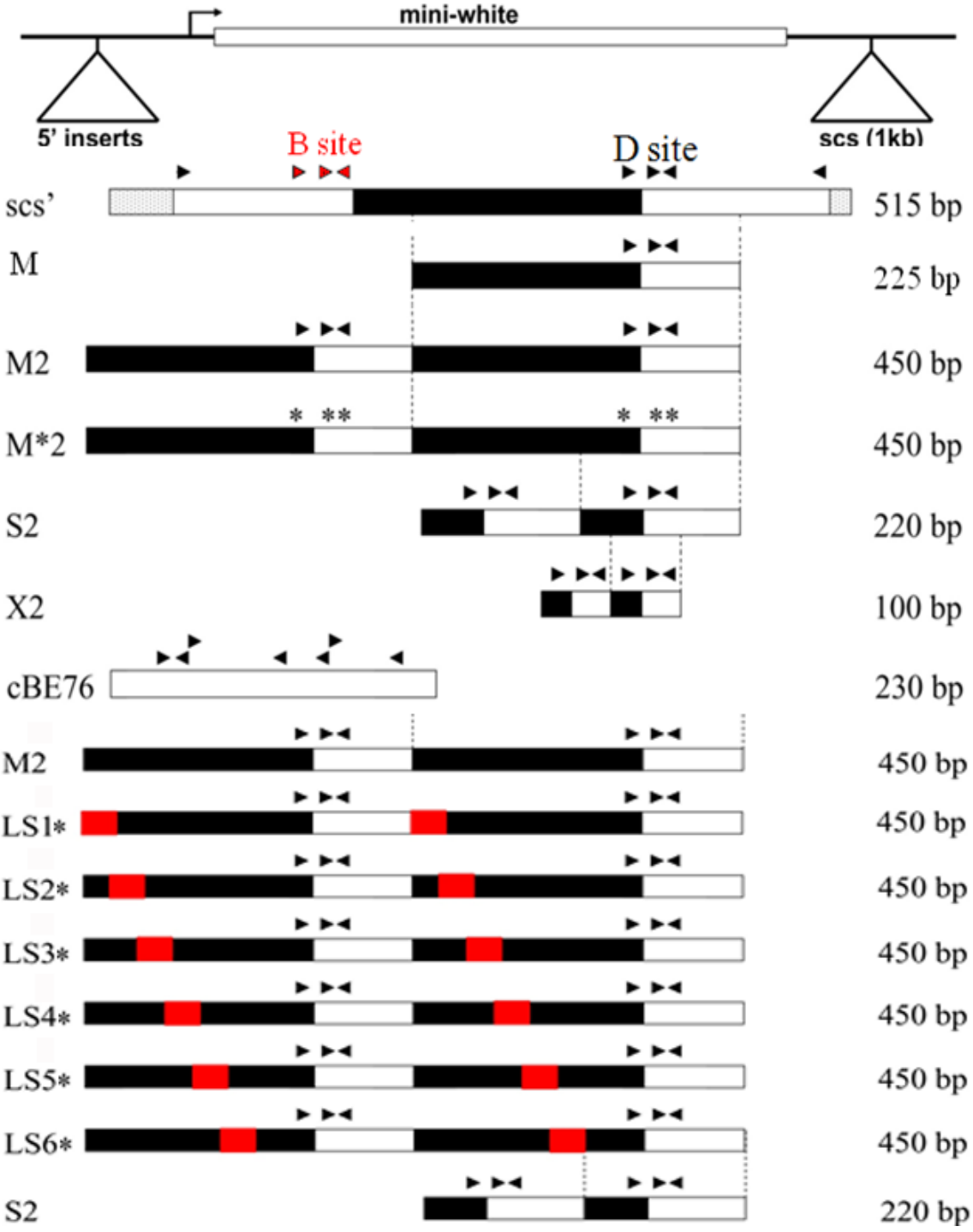

Figure 2.3 Subfragments of scs' used in position effect assays to detect insulator activity. Arrowhead represent CGATA motif, star (*) in M*2 fragment indicate mutated CGATA motif. Red boxes are mutations of LS1 to LS4, named as LS1* to LS6*. cBE76 is a genomic DNA sequence containing BEAF binding motif.

Table 2.1 Scs' derivatives in position effect assays

5'Insulator	Eye color (% of lines)				5'Insulator	Eye color (% of lines)			
	Light	Dark	No.	Activity		Light	Dark	No.	Activity
-/-	34	66	12	-	S2	72	28	21	+
-	50	50	38	-	X2	72	28	25	+
scs'	90	10	10	++	LS1*	87	13	39	++
M2	90	10	20	++	LS2*	87	13	63	++
M*2	30	70	10	-	LS3*	100	0	12	++
cBE76	100	0	12	++	LS4*	67	33	24	+
					LS5*	100	0	10	++
					LS6*	91	9	23	++

In all cases except -/-, scs is placed downstream of the *mini-white* gene; 5' insulator represents the insulator upstream of *mini-white*. Star (*) represents mutation. Kolmogorov Smirnov two sample test was used to measure statistical significance of insulator activity. Light eye color includes yellow, orange; dark eye color includes dark orange, light red and red.

We divided the M fragment into five subfragments: 5' end, LS4 region, spacer region, D site, and 3' end. M5Δ, M3Δ and M35Δ fragments were made from the M fragment by deleting the 5' end, the 3' end, or both the 5' and 3' ends, respectively. M35SΔ fragment is made by deleting 5' end, 3' end and the spacer region from the M fragment. M35ΔS* fragment is made by deletion of 5' and 3' ends, and by mutation in the spacer sequences (Figure 2.4). These fragments were cloned into the pC4 *attB* plasmid for fly injection to detect insulator activity in position effect assays. Here, we used the *Φ*C31integrase mediated site specific insertion system. In this system, *Φ*C31integrase mediates specific integration of an *attB* site into a single transgenic *attP* landing site already in the fly genome.

The P element is a *Drosophila* transposon that has been widely used in *Drosophila* genetic research (Rubin and Spradling, 1982; Spradling and Rubin, 1982). One of the main features of P element transformation is that it inserts randomly into the *Drosophila* genome. Because of the random insertion, 10-20 transgenic fly lines should be obtained in order to make a statistically convincing conclusion about whether or not a test sequence functions as an insulator. This is time and energy consuming. Additionally, the efficiency of transforming large DNA fragments

into the *Drosophila* genome by P elements is low. FLP/*FTR* and Cre/*LoxP* site specific recombination systems have been applied in *Drosophila* to reduce the considerable work caused by random insertion (Horn and Handler, 2005; Oberstein et al., 2005). Pairs of transgenes flanked by *FRT* and *LoxP* sites could be placed at the same position in genome by P element transformation (Siegal and Hartl, 1996). However, these techniques are often used in applications, such as mosaic assays, rather than site specific integration. Therefore, they are not suited for precise targeted transgene integration.

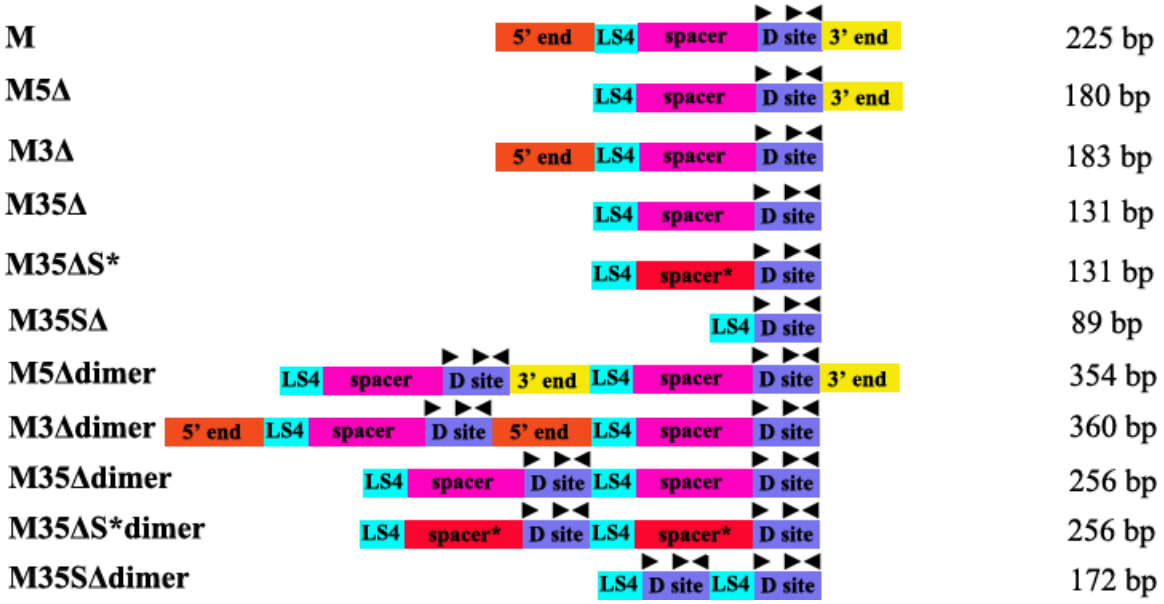

Figure 2.4 Candidates of minimal scs' insulator sequences. Blue, aqua, pink, red, orange, yellow represents BEAF binding D site, LS4 region, spacer between LS4 and D site, mutated spacer between LS4 and D site, 5' end and 3' of M fragment, respectively.

The *Φ*C31 integrase system provides a novel strategy for germline transformation (Bischof et al., 2007; Venken et al., 2006) (Figure 2.5). The integrase encoded by bacteriophage *Φ*C31 permits the sequence specific integration between a bacterial attachment site (*attB*) and a phage attachment site (*attP*) to create *attL* and *attR* sites. Unlike FLP and Cre recombinases, *Φ*C31 integrase only mediates integration. A modified form of the integrase mediates the recombination between *attL* and *attR* sites to form *attB* and *attP* sites; this is called disintegrase

(Figure 2.5) (Thorpe et al., 2000). The advantages of the *Φ*C31 system are *Φ*C31 integrase mediates precise targeting into *attP* landing sites which are engineered into *Drosophila* chromosomes. Different sequences can be tested at the same location, eliminating variable chromosomal position effects, so fewer fly lines are needed. Second, constitutive expression of *Φ*C31 integrase in *Drosophila* does not affect chromosomal stability, so fly lines with endogenous *Φ*C31 integrase gene have been generated. Coinjection of *Φ*C31 mRNA or a helper plasmid containing the *Φ*C31 gene is not necessary when doing germline transformation. Third, large DNA fragment transformation may be easier.

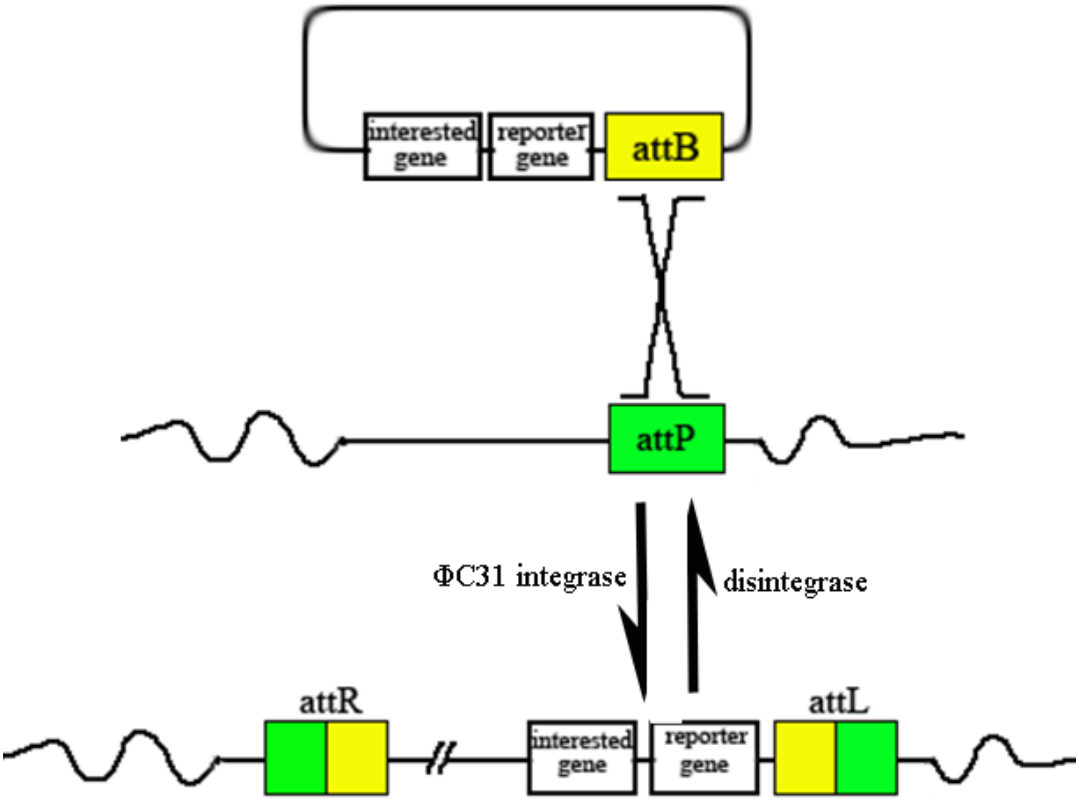

Figure 2.5 Strategy of *Φ*C31 integrase mediated site specific integration. *Φ*C31 integrase catalyze the recombination between *attB* site from bacteria and *attP* site in drosophila genome, whereas engineering modified disintegrase mediate the recombination between *attR* and *attL*.

First, we tested 13 *attP* fly lines from the Bloomington Drosophila Stock Center (BDSC) by injecting pC4-*attB* and pC4-M2-*attB* plasmids. Four lines showed weak position effects, six did

not show position effects, and three lines were unhealthy. Therefore, we are making flies with new *attP* landing sites for our purposes.

2.2 MATERIALS AND METHODS

2.2.1 pSyn Series Constructs

M5Δ fragment was PCR amplified from pSyn-M plasmid using scs'-M-5'del-5'-*Bgl*II primer and scs'-M-5'del-3'-*Bam*HI primer. M3Δ fragment was PCR amplified from pSyn-M plasmid using scs'-M-3'del-5'-*Bgl*II primer and scs'-M-3'del-3'-*Bam*HI primer. M35Δ fragment was PCR amplified from pSyn-M plasmid using scs'-M-5'del-5'-*Bgl*II primer and scs'-M-3'del-3'-*Bam*HI primer. M35spacerΔ fragment was generated by annealing scs'-LS4-D-5'-*Bgl*II primer and scs'-LS4-D-3'-*Bam*HI primer in 94 ℃ H_2O. M35spacer* fragment was generated by annealing scs'-LS4-spacer-D-5'-*Bgl*II primer and scs'-LS4-D-3'-*Bam*HI primer in 94 ℃ H_2O. M3Δ, M5Δ, M35Δ, M35spacerΔ, M35Δspacer* monomer fragments were cloned in pSyn vector which was isolated from *Bam*HI/*Bgl*II digestion of pSyn-M plasmid. In order to get pSyn-M derivative dimer plasmids, all of pSyn series monomer plasmids were cut with *Sca*I/*Bam*HI and *Sca*I/*Bgl*II, respectively. Two of the fragments with *Sca*I/*Bam*HI end and *Sca*I/*Bgl*II end were ligated to generate pSyn series M derivatives dimer plasmids. pSyn-M* (M fragment with BEAF binding site D site mutation), pSyn-M*2 (two adjacent M fragments with two mutated D sites) and PSyn-LS4*(M fragment with LS4 region mutation) plasmids were made by previous labmates. In order to generate a pSyn-M*LS4* which possesses D site mutation and LS4 region mutation, LS4 mutation site was introduce into pSyn-M* plasmids by Quickchange Site-directed Mutagenesis Kit (Stratagene) with LS4-52' primer and LS4-32' primer. New pSyn-scs' series plasmids were sequenced with syn-M-5'-27 primer and Syn-M-new-3' primer.

2.2.2 pC4-scs-*attB* Series Constructs

M derivative fragments, cut with *Bam*HI/*Bgl*II, were cloned in pC4-scs-*attB* vector which was cut by *Bam*HI. Generated Plasmids in which *Bam*HI sticky end of M derivatives was ligated with *Bam*HI sticky end of the vector were selected as the injection plasmids. The C4scs-5' primer and C4scs- new-3' primer were used to sequence new pC4-scs-*attB* plasmids. The pC4-scs-*attB* plasmids we made are in table 2.2.

Table 2.2 pC4-scs-*attB* plasmids for injection

AttB constructs monomer	*AttB* constructs dimer
pC4-scs- M-*attB*	pC4-scs- M2-*attB*
pC4-scs- M*-*attB*	pC4-scs- M*2-*attB*
pC4-scs- M3Δ-*attB*	pC4-scs- M3ΔD-*attB*
pC4-scs- M5Δ-*attB*	pC4-scs- M5ΔD-*attB*
pC4-scs- M35Δ-*attB*	pC4-scs- M35ΔD-*attB*
pC4-scs- M35spacerΔ-*attB*	pC4-scs- M35spacerΔD-*attB*
pC4-scs- M35Δspacer*-*attB*	pC4-scs- M35Δspacer*D-*attB*

2.2.3 pRLY Constructs

M2 fragment, first digested with *Bam*HI/*Eco*RI then filled in the sticky end by Klenow, was inserted between two FRT sites in pUC-FNF plasmid. Double stranded *attR* oligonucleotides were generated by mixing attR-5b' primer and attR-3b' primer in 95 ℃ H_2O until they were cooled down to room temperature. We cloned double stranded *attR* oligonucleotides into pUC-FNF-M2 vector digested with *Sph*I/HindIII. Double stranded *attL* oligonucleotides were generated by mixing attL-*Nsi*I-5' primer and attL-*Sph*I-3' primer in 95 ℃ H_2O until they were cooled down to room temperature. The small fragment, cut by *Nde*I was self-ligated to make a

circular plasmid named p-scs. Double stranded *attL* containing *Nsi*I sticky end was cloned in the new p-scs plamid which was cut by PstI. The new p-scs-attL plasmid was religated with other half of pC4-scs plasmid to makepC4-scs-attL plasmid. New attR-FRT-M2-FRT fragment, digested by *Xho*I/*Eco*RI, was cloned in the pC4-scs-attL vector which was cut by *Xho*I/*Eco*RI. This new plasmid was called pC4-scs-attL-FRT-M2-FRT-attR. The *yellow* gene was PCR amplified from pCaspeR4 plasmid with yel-*sph*-5' primer and yel-*sph*-3' primer. The *yellow* gene was inserted into new pC4-scs-attL-FRT-M2-FRT-attR by *Sph*I digestion. This final plasmid containing pC4-scs-attL-FRT-M2-FRT-attR-yellow was called pRLY plasmid.

2.2.4 Fly Stocks and Germline Transformation

Flies were maintained on standard cornmeal, yeast, and dextrose with Tegosept at 25 ℃. $y^{-}w^{-}$ flies or *attP*/integrase flies were used for plasmid injection. Disintegrase flies were kind gifts from Dr. Fran ̧ois Karch (University of Geneva). Lines from Bloomington Drosophila Stock center were ZH22A (24481), ZH51C (24482), ZH51D (24483), ZH58A (24484), ZH68E (24485), ZH86Fa (24486), ZH86Fb (24749), ZH96E (24487), ZH102D (24488), VK00020 (24867), VK00031 (24870), VK00033 (24871), VK00037 (24872), delta 2-3 transposase fly and flippase fly. The pRLY transgenic flies were generated by co-injection of plasmids (0.4 μg/μL) and the pπ25.7wc helper plasmid (0.14 μg/μL) into $y^{-}w^{-}$ embryos. pC4 and pC4-M2 flies were generated by injection of pC4-scs-*attB* and pC4-scs-M2-*attB* constructs (0.4 μg/μL) into *attP*/integrase flies.

2.2.5 Generation of New *AttP* Landing Site Flies

Transgenic flies generated by P element transformation were crossed with $y^{-}w^{-}$ flies and, Cyo/Sp^{1} flies and $TM3/Scm^{ET50}$ flies, respectively, to determine which chromosome the transgene is on. After the transgene landing chromosome is established, transgenic flies are crossed with

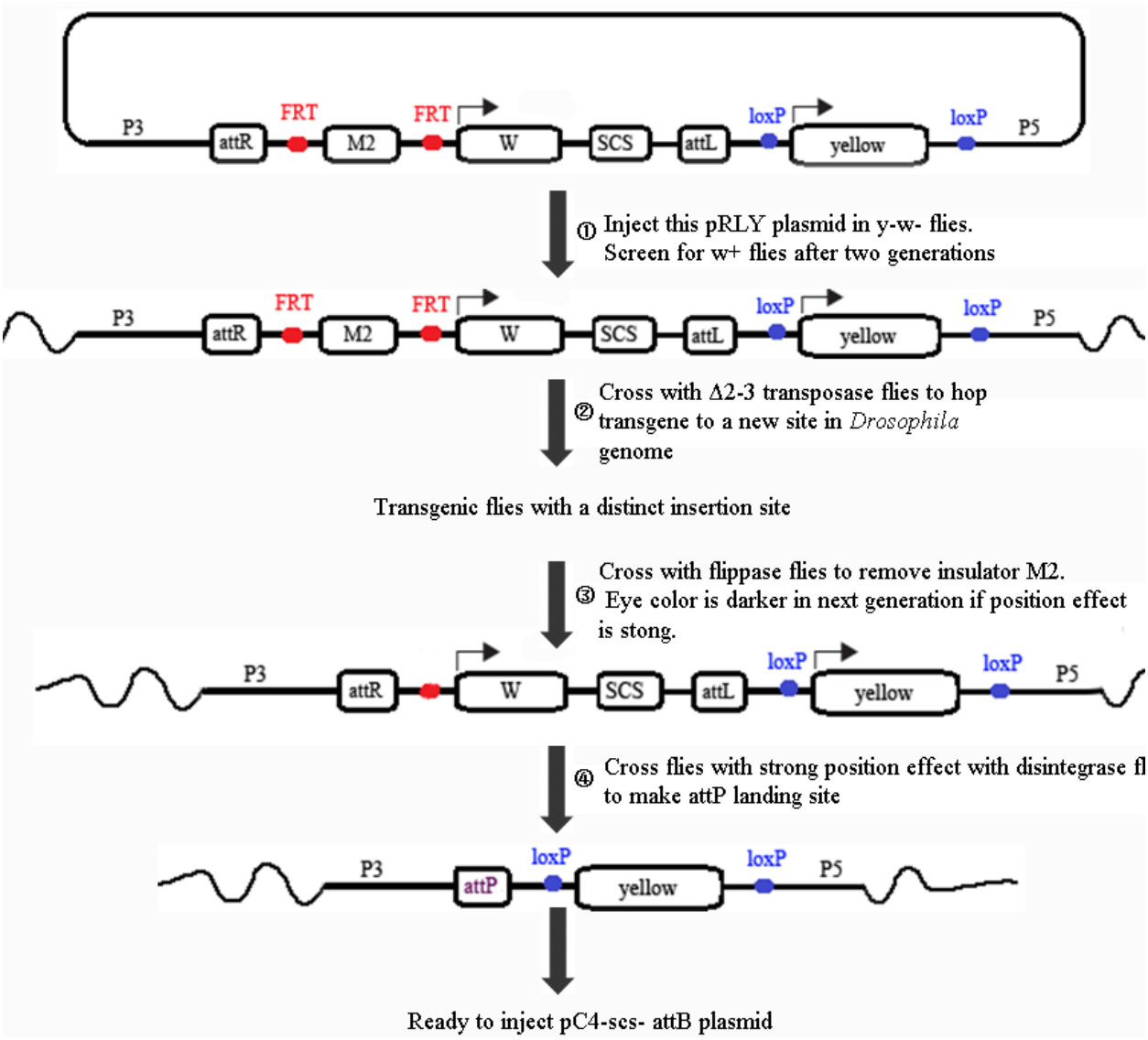

Figure 2.6 Diagram of making *attP* landing site flies and hopping around transgene. Red and blue ovals represent FRT and LoxP sites, respectively. Step ② is to hop around transgene to obtain more insertion sites.

delta 2-3 transposase flies to induce transposition around the *Drosophila* genome to obtain new transgene landing sites. New transgenic flies were crossed with flippase flies. The embryos of their offspring are incubated at 37 ℃ for 1h to heat shock the promoter of flippase. In this generation, screen flies with darker eye colors which indicate that the insulator is removed. Single fly genomic PCR was performed to confirm insulator was removed. More details of strategy to make *attP* flies are in Figure 2.6.

2.2.6 Single Fly DNA Preparation and PCR

One male or female fly was placed in a 1.5 ml tube and mashed for 5-10 seconds with a pipette tip containing 50 µL squishing buffer (SB) (10 mM Tris pH 8.2, 25 mM NaCl, 1mM EDTA, 0.2 µg/µL Proteinase K). Incubate at 37 ℃ for 30 min and then 95 ℃ for 10 min. Spin down for 1 min at the highest speed and place the tube on ice. The fly genomic DNA is ready for PCR use.

In 20 µL reaction, 1.5 µL single fly DNA supernatant was used as template. 200 nM P3-(delM2/attRL)-5' and wh-prom-(delM2)-3' were used as primers. First, denature for 1 min at 94 ℃. Second, denature at 94 ℃ for 5 sec, anneal at 52.9 ℃ for 15 sec, and extend at 72 ℃ for 62 sec. This cycle was repeated 25 times. Third, further elongate for 10 min, and hold overnight at 10 ℃ if necessary.

2.3 RESULTS

Out of 12 *attP* fly lines (VK lines and ZH lines) tested from the Bloomington Drosophila Stock Center (Bischof et al., 2007; Venken et al., 2006), one VK line (VK00020) and three ZH lines (86Fa, 58A and 102D) showed chromosomal position effects that were blocked by M2, based on lighter eye color (Figure 2.7). Three lines were not healthy and six lines did not show a position effect (Table 2.3).

Three transgenic fly lines F3F1, F7F1 and F7F2 were generated with the pRLY construct by P element transformation. F7F1 and F7F2 were from the same injected fly. All fly lines have yellow eye color and dark body color (Figure 2.8). Crossing transgenic flies with balancer chromosomes determined that pRLY is on the X chromosome in the F3F1line and on the 2nd chromosome of the F7F1and F7F2 fly lines. After crossing with flippase flies to remove the M2 fragment, unfortunately, there was no apparent difference of eye color (Figure 2.8A). One fly

Table 2.3 Summary of transgenic *attP* fly lines obtained from BDSC

Fly lines	Insertion site	Eye color		Position effect	Viability
		pC4-scs-*attB*	pC4-scs-M2-*attB*		
ZH22A	22A2	Unknown	Unknown	Unknown	Unhealthy
ZH51C	51C1	Yellow	Yellow	_	Viable
ZH51D	51D9	Dark orange	Dark orange	_	Viable
ZH58A	58A3	Orange	Yellow	+	Viable
ZH68E	68E1	Orange	Unknown	Unknown	Unhealthy
ZH86Fa	86E18	Dark orange	Orange	+	Viable
ZH86Fb	86F8	Orange	Orange	_	Viable
ZH96E	96E10	Unknown	Red	Unknown	Unhealthy
ZH102D	102D	Red	Light red	+	Viable
VK00020	99F8	Yellow	Light yellow	+	Viable
VK00031	62E1	Yellow	Yellow	_	Viable
VK00033	65B2	Yellow	Yellow	_	Viable
VK00037	22A3	Yellow	Yellow	_	Viable

Both ZH lines and VK lines are *attP* fly lines and obtained from BDSC. Two constructs pC4-scs-*attB* and pC4-scs-M2-*attB* were tested to see chromosomal position effect. Three lines ZH22A, ZH68E and ZH98E were unhealthy. Eye color was taken pictures for two-day old female heterozygous for indicated transgene.

PCR confirmed that the M2 fragment was removed from the transgene. The flippase recombination efficiency is high, 9 out of 10 flies lost M2 fragment after crossing with flippase flies and heat shocking for 1h at 37 ℃ (Figure 2.9).

In order to get more transgene landing sites, these fly lines were crossed with delta2-3 transposase flies. The strategy for transposing the element is described in Figure 2.6. In the 2nd generation, it can be determined if the M2 transgene has moved to new insertion sites. Out of 37

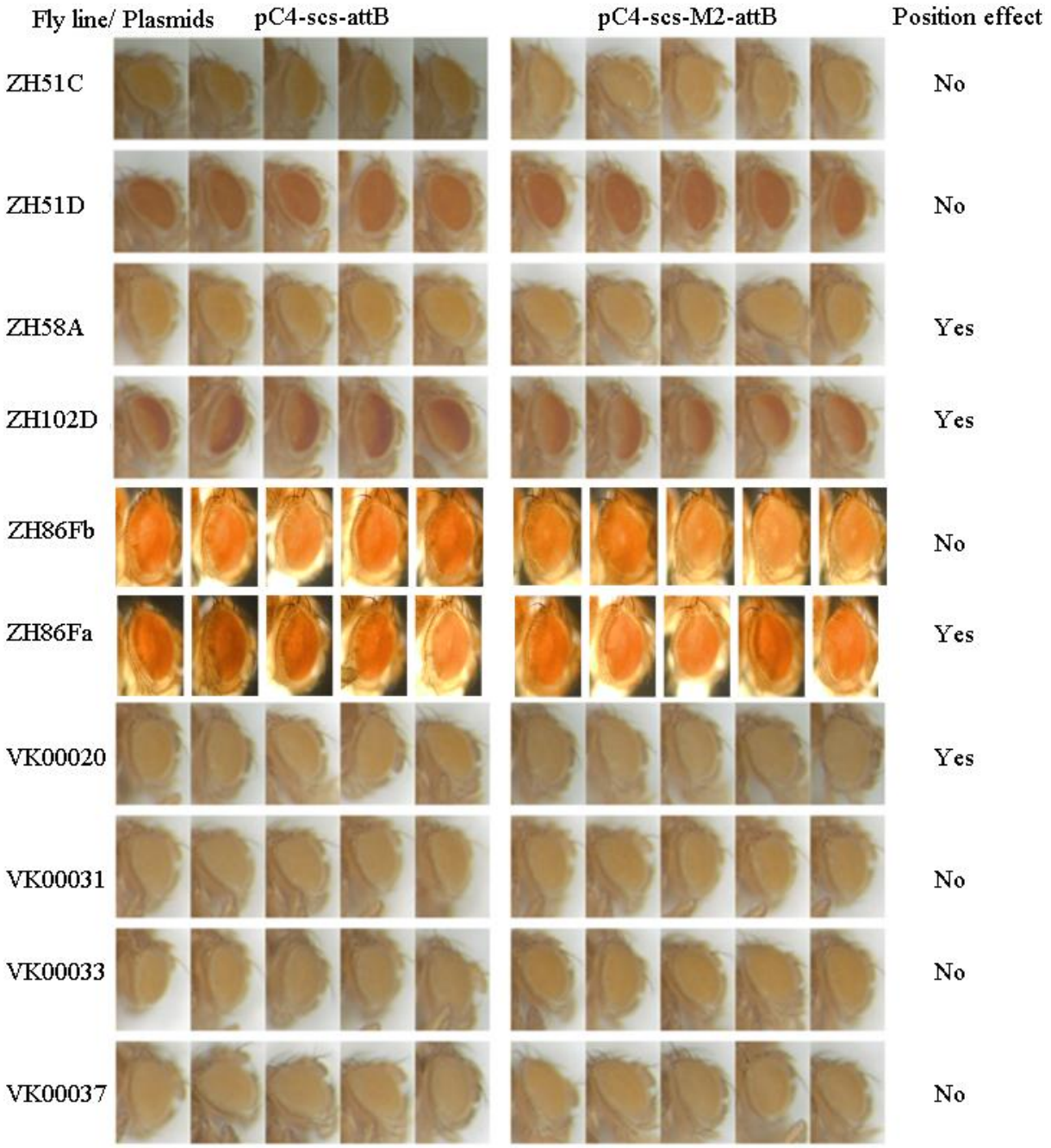

Figure 2.7 Position effect assays in available *attP* fly lines. pC4-scs-*attB* and pC4-scs-M2-*attB* plasmids were injected in *attP* fly lines. The *attP* landing sites of ZH51C, ZH5 1D and ZH58A lines are on the second chromosome. The *attP* landing site of 102D is on the fourth chromosome. VK00020, VK00031, VK00033 and VK00037 are VK lines. VK00020, VK00031 and VK00033 have *attP* landing sites on the third chromosome, whereas VK00037 have on second chromosome. Eyes of 2-day old female heterozygous for all indicated transgene are shown.

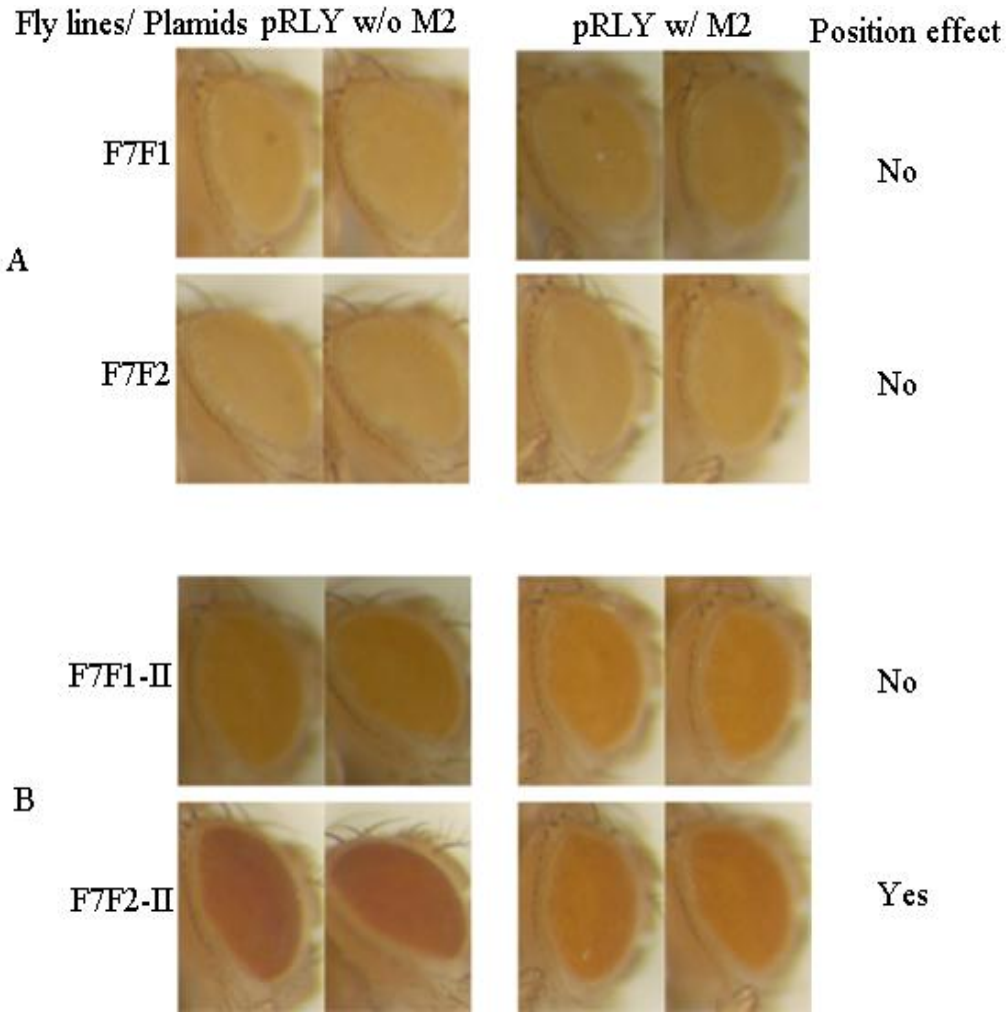

Figure 2.8 Position effect assays of flies with pRLY construct. (A) F7F1 and F7F2 are two flies from injection with pRLY plasmid. (B) F7F1-I and F7F2-I are transgenic flies whose transgene were hopped to a new site from F7F1 and F7F2, respectively. All of the eyes are from 2-day old heterozygous transgenic females.

transposition crosses of F7F1 flies, only 11 crosses were found with transpositions, suggesting that the transposition efficiency of this line is around 30%. Out of 46 crosses with the F7F2 flies, transpositions were found in 15 vials, giving an efficiency of around 33%. Note that there might be more than one transposition event occurred in each vial. Unfortunately, out of 47 crosses of F3F1 with transposase flies, no transposition flies were found in any vial. Thus, the transgene in the F3F1 chromosome appears to be unable to hop, possibly because one or both ends of the P element are not intact. It was determined that three crosses induced transposition of the transgene

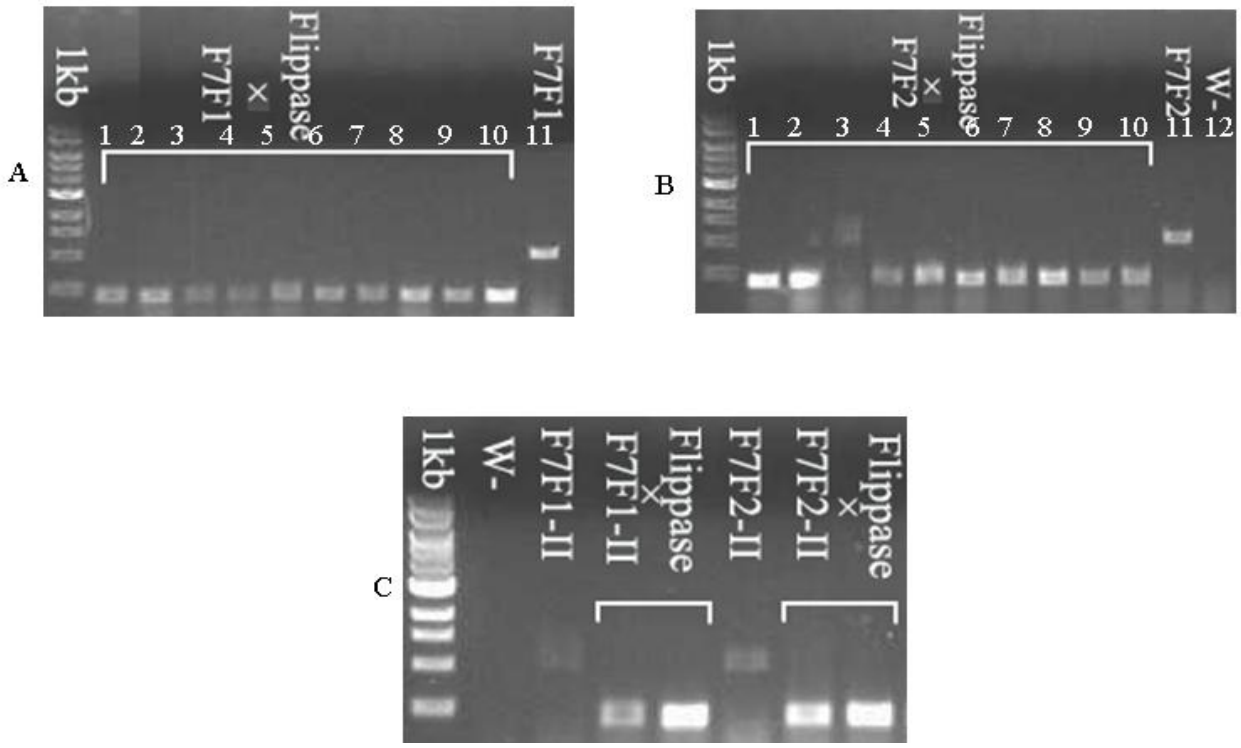

Figure 2.9 Single fly PCR of pRLY flies. (A) F7F1 flies w/ and w/o M2 insulator. Lane 1 to lane 10 are single flies from different crosses with flippase. Lane 11 is fly with M2 insulator as a control. (B) F7F2 flies w/ and w/o M2 insulator. Lane 1 to 10 are single flies from different crosses with flippase. Lane 11 is fly with M2 insulator, lane 12 are w^- flies without transgene as a negative control. (C) F7F1-II and F7F2-II w/ and w/o M2 insulators. w^- is a negative control.

onto the *CyO* chromosome, eight hopped the transgene onto the *Sb* chromosome, four hopped onto the wild type second chromosome, and six hopped onto the wild type third chromosome. Then the transgene on the *CyO* or *Sb* chromosomes were hopped again, unfortunately, only one line was able to hop. Three more fly lines with transgene on wild type second or third chromosome were generated.

In total, I obtained 15 new pRLY fly lines with transgene on wild type second or third chromosome. Three show strong position effects. Of these lines, one is homozygous lethal, one is homozygous sterile, and the third is homozygous lethal with the eye color change reversed, suggesting a silencing chromosomal position effect (Table 2.4).

Table 2.4 Summary of transgenic pRLY fly lines

Fly lines	Insertion chromosome	Eye color		Position effect	Viability
		pRLY w/o M2	pRLYw/ M2		
F3F1	1	Unknown	Orange	_	Lethal
F7F1	2	Yellow	Yellow	_	Lethal
F7F1-II	3	Orange	Orange	_	Lethal
F7F1-III	3	Orange	Orange	_	Lethal
F7F1-IV	3	Orange	Orange	_	Lethal
F7F2	2	Yellow	Yellow	_	Lethal
F7F2-II	3	Red	Orange	+	Lethal
F7F2-III	3	Yellow	Orange	+ (silencing)	Lethal
F7F2-IV	3	Light red	Red	+	Sterile
F7F2-V	3	Light red	Light red	_	Viable
F7F2-VI	3	Light red	Light red	_	Viable
F7F2-VII	3	Light red	Light red	_	Lethal
F7F2-VIII	2	Orange	Orange	_	Lethal
F7F2- IX	2	Dark orange	Dark orange	_	Lethal
F7F2-XI	3	Orange	Orange	_	Lethal

F7F1, F7F2 and F3F1 were the original transgenic flies obtained from injection. Other fly lines were obtained through crossing with Δ2-3 transposase flies. F7F2-III flies eye color was much lighter when M2 insulator was removed, indicating the transgene was located in a silenced domain.

2.4 DISCUSSION

It is surprising that most of the fly lines do not show a position effect. Previous results suggest that at least 50% of insertion sites show chromosomal position effects. Presumably most chromosomal positions sampled in this study lack nearby activating regulatory elements. The reason for the difference compared to earlier studies (Table 2.1) is not clear. Fly lines that show position effects are recessive lethal or show weak position effects. Transposition is still in progress to obtain additional insertion sites in order to get viable flies with strong position effects.

It is not clear why the flies that have undergone transposition are homozygous lethal or sterile. This is unusual.

2.5 FUTURE WORK

Continue performing hopping crosses until at least two lines are found that have strong position effects. After obtaining two working fly lines, crosses to disintegrase flies will make new *attP* landing site flies. Then the various M derivatives will be tested for insulator function in these flies to identify minimal scs' sequence for insulator activity. Meanwhile, RT-PCR should be performed to see if insulator activity is related with promoter activity.

CHAPTER 3
DETECTION OF SCS' PROMOTER ACTIVITY TO SEE HOW PROMOTER ACTIVITY CORRELATES WITH INSULATOR ACTIVITY

3.1 INTRODUCTION

Recently, through ChIP-chip our lab found there are at least 1820 BEAF binding sites in the *Drosophila* genome. More than 85% of these binding sites are located within 300 bp of transcription start sites (TSSs). Half of the binding sites are between head-to-head divergent gene pairs (Jiang et al., 2009). For example, the scs' insulator has two divergent promoters: the *aurora* and *CG3281* promoters. Most BEAF-associated genes are highly expressed in diverse tissues (Bushey et al., 2009; Jiang et al., 2009; Negre et al., 2010). Intrigued by this, our lab decided to identify the relation of BEAF and these genes. Our lab performed RT-PCR to detect the expression level of BEAF associated genes in embryos with a null mutation in BEAF. The expression level of most tested genes dropped in the absence of BEAF (Jiang et al., 2009). Similar results were obtained after RNAi knockdown of BEAF in cultured cells. These results indicate that BEAF may play a role in maintaining the high expression of BEAF-associated genes.

Here, we present evidence that the M fragment does function as a promoter in transfected cells. The M fragment mutations described in the previous chapter will be tested for promoter activity to see if it correlates with insulator activity.

3.2 MATERIALS AND METHODS

3.2.1 DNA Constructs

Two constructs pTRW-fibrillarin containing mRFP-fibrillarin fusion gene and pTGW-Nop56 containing EGFP-Nop56 fusion gene were kind gifts from Dr. DiMario (Louisiana State University). Both mRFP-fibrillarin fusion gene and EGFP-Nop56 fusion gene were PCR amplified with TRW- *Sal*I-3'and TRW-*Xba*I-5'. The PCR was performed under for 25 cycles

under following conditions: 98 ℃ for 5 sec, 55 ℃ for 15 sec and 72 ℃ for 45 sec. mRFP-fibrillarin, containing a high G-C rich region, was amplified with Phusion High Fidelity PCR Kit (Biolabs, Inc) according to the manufacture. The vectors were made by digesting of pSyn-M with *Sal*I and *Xba*I. pSyn- M-mRFP-fibrillarin and pSyn-M-EGFP-Nop56 constructs were generated by ligating *Sal*I/*Xba*I cut vectors and PCR products with *Sal*I/*Xba*I end. Other M derivative fragments- M*, LS4*, M*LS4*, M3Δ, M5Δ, M35Δ, M35spacerΔ and M35Δspacer*- were attached with EGFP-Nop56, the procedures were the same as M fragment. Control plasmid pEK-CMV-GFP-Nopp140 containing cytomegalovirus promoter was obtained from Dr. DiMario. Midi-preps were done for control plasmid and all of M and M derivative plasmids with reporter fusion gene (Quiagen). Actin5C promoter isolated from *Bam*HI/*Eco*RI digestion of pPac plasmid was cloned into EGFP-Nop56 vector isolated from *Bam*HI/*Eco*RI digestion of pSyn-M-EGFP-Nop56 construct. This new control plasmid with actin5C promoter is named pSyn-actin5C-EGFP-Nop56.

3.2.2 Cell Culture and Transient Transfection

Drosophila Schneider 2 (S2) cells are provided by Dr. DiMario (Louisiana State University) and grown at 25 °in M3 (Shields and Sang M3 insect media, Sigma) media plus 10% FBS (fetal bovine serum, Invitrogen) and 100 units/ml Pen/Strep (Penicillin/Streptomycin, Invitrogen). S2 cells are split every week at a 5 times dilution. Calcium phosphate transfection was performed according to Kingston, R. E (Kingston, 2003). Lipofectamine LTX and PLUS Reagents Kit (Invitrogen) was used based on the protocol described by the manufacture. Both Calcium phosphate and lipofectine methods were either inconsistent or inefficient. Effectene transfection reagent (Quiagen) was used for S2 cell transient transfection in 12 well plates. 0.4×10^5/ml S2 cells were grown on 12 well plates the day before transfection. 0.2 μg DNA, 1.6 μL enhancer

(ratio of DNA: enhancer is always 1:8) and 5 μL effectene were used for each well. Vortex for 10 sec immediately after adding effectene, followed by incubation at room temperature for 15 min. Leave the cell in the complete media with antibiotics containing transfection reagent for five days. Take out cells and check if fluorescent proteins were produced with fluorescence microscopy (SPOT Imaging Solutions).

3.2.3 Western Blot

0.4×10^5/ml transient transfected 4 days S2 cells were mixed with SDS sample buffer (125 μM Tris/SDS containing 0.1% SDS pH 6.8, 10% glycerol, 0.2% SDS, 100 μM DTT and 1.5 nM bromphenol blue). Primary rabbit anti-GFP antibody was used at a 1:5000 dilution, primary mouse anti-fibrillarin antibody was used at a 1:2 dilution. The secondary goat anti-rabbit and goat anti-mouse antibody conjugated with HRP were used at a 1:10000 dilution, respectively.

3.3 RESULTS

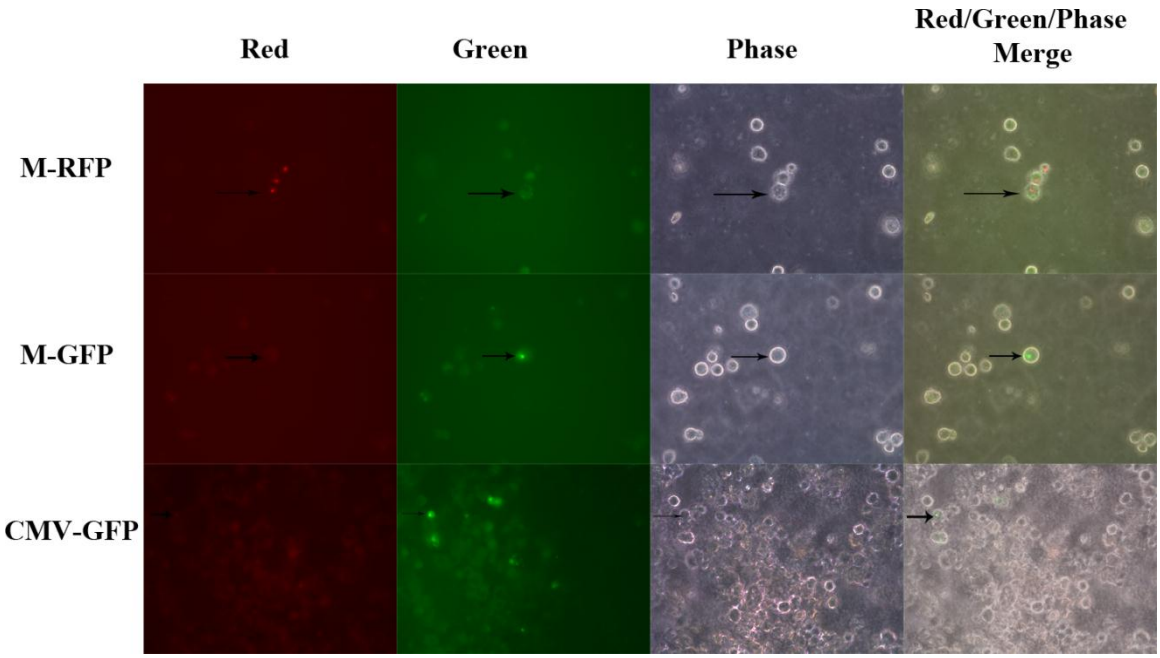

Figure 3.1 Transient transfection of S2 cells. S2 cells were maintained in lipofectin transfection reagent for five days. pSyn-M-mRFP-fibrillarin and pSyn- M-EGFP-Nop56 were tested for promoter activity. pEK-CMV-GFP-Nopp140 plasmid as a positive control. Arrowhead points to transfected S2 cells.

Currently only control and M fragment plasmids were tested. In figure 3.1, Red and green fluorescence were only seen in pSyn-M-RFP-fibrillarin and pSyn-M-GFP-Nop56, respectively, indicating M functions as a promoter. However, because the transfection efficiency was very low,

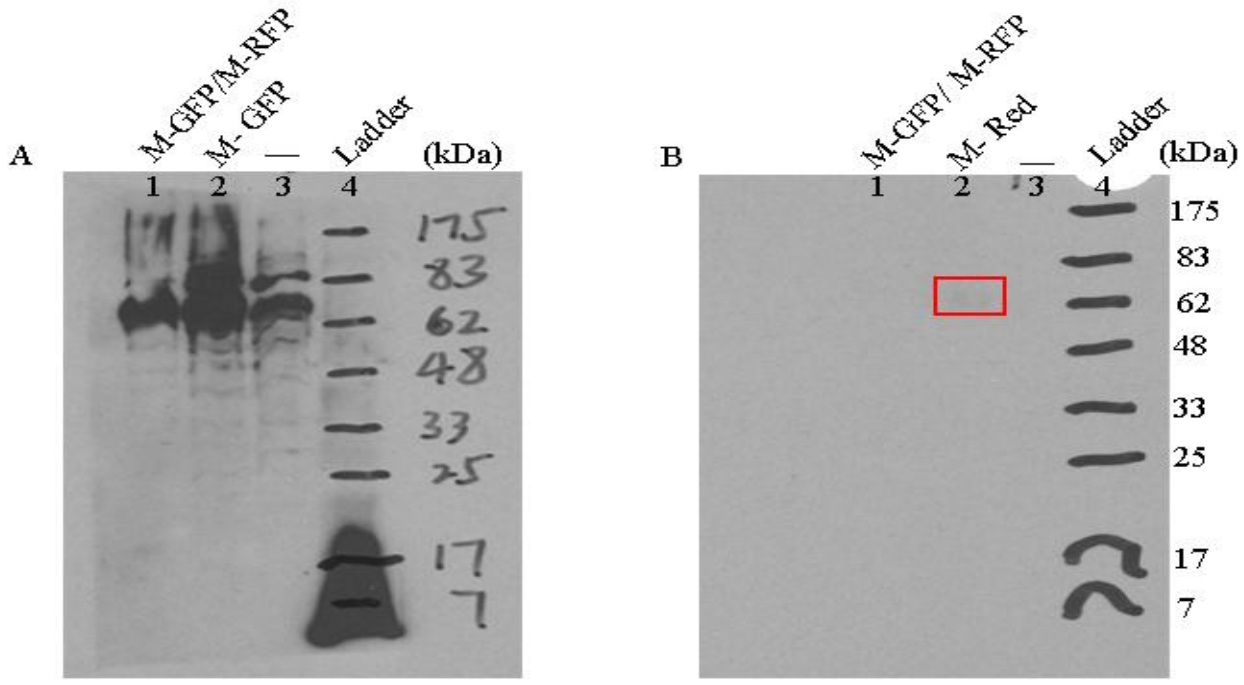

Figure 3.2 Western blot analysis of transient transfected S2 cells. (A) M- GFP represents single transfection of PSyn- M- GFP- Nop56. (B) M- Red stands for single transfection of PSyn- M- RFP- fibrillarin. Faint band in the red box was RFP- fibrillarin fusion protein. M- GFP/ M- RFP in both (A) and (B) represent transient cotransfection of PSyn- M-GFP- Nop56 and PSyn- M- RFP- fibrillarin into S2 cells. Minus represents a negative control in which no transfection was applied in S2 cells.

it was hard to tell if the M promoter was stronger than the control CMV promoter or not.

Western blot analysis was inconclusive. There were protein bands in all lanes using anti-GFP antibody, even in untransfected control S2 cells. The fusion protein of EGFP-Nop56 should be 85 kDa (Figure 3.2A). For the pSyn-M-RFP-fibrillarin transfection, the endogenous fibrillarin and the fusion protein of RFP-fibrillarin should be 34 kDa and 62 kDa, respectively. A very faint band around 62 kDa was seen in M-RFP-fibrillarin lane (Figure 3.2B). Surprisingly, we didn't detect fibrillarin in double plasmid transfection lane. Endogenous fibrillarin was not detected

either. Other M mutation plasmids were tested by transient transfection. The results were not conclusive because of low transfection efficiency and photobleach (data not shown).

3.4 DISCUSSION

Although we did observe M has promoter activity, two problems bothered us. One is the transfection efficiency was too low and inconsistent, at best 10%. Second problem is photobleaching happened too fast to take a picture. In order to do avoid photobleaching, we came up with chemical methods like western blot to test fluorescent proteins. In order to get a decent amount of fluorescent proteins which could be detected with western blot, stable transfection is recommended to accumulate fluorescent proteins. Additionally, control pEK-CMV-GFP-Nopp140 was often expressed at a low level. It was reported that cytomegalovirus promoter is expressed very low in *Drosophila* S2 cells. We made a new control plasmid pSyn-actin5C-EGFP-Nop56 containing actin5C to overcome this difficulty (Qin et al., 2010).

3.5 FUTURE WORK

Since the transfection efficiency was so low, stable transfection may be a good strategy to select and enrich for cells that take up DNA. pSyn-M-RFP-fibrillarin, pSyn-actin5C-EGFP-Nop56 plasmids and pNeo plasmid containing neomycin resistant gene are cotransfected into S2 cells. After 72 hours, selective media is added with the concentration of 1 μg/ml G418 (Sigma) to select cells with neomycin resistant gene integrating into the genome. Cells are split and selective media is replaced every 5-7 days. At 3 weeks, cells are visualized and images are captured using fluorescence microscope. Meanwhile, western blot against GFP and fibrillarin could be applied to detect the fused fluorescent protein level. If stable transfection works, other M derivative plasmids could be done in the same way to answer the question if the BEAF binding site, LS4 sequence or both are necessary for promoter activity.

CHAPTER 4
PURIFICATION OF BEAF ASSOCIATED PROTEINS

4.1 INTRODUCTION

The original electrophoretic mobility shift assay (EMSA) with scs' detected binding by proteins that resulted in a low mobility complex, in addition to binding by BEAF (Zhao et al., 1995). Other experiments found that while the high affinity BEAF binding site in scs' is essential for insulator activity, other sequences contribute to insulator efficiency. In particular, mutating the LS4 sequences weakens the insulator activity. This suggests that proteins bind to LS4, perhaps the proteins that make the low mobility complex with the scs' sequence.

Here we used EMSA to detect LS4 binding proteins. Further characterization needs to be done to investigate sequence specificity for LS4, in particular, DNase I footprinting. We used an oligonucleotide with the LS4 and D site sequences to affinity purify proteins. If specificity for LS4 can be demonstrated, the proteins can be identified by mass spectrometry.

4.2 MATERIALS AND METHODS

4.2.1 Cell Culture and Preparation of Kc Cell Nuclear Extract

Drosophila Kc167 cells were obtained from Drosophila Genomics Resources Center (stock #:1), and they are maintained at 25℃ in M3 media containing 5% FBS, 100 units/ml Pen/Strep and BPYE (Bacto-peptone and yeast extract, USBiological). The cells were split every week at a 10 times dilution. *Drosophila* S2 cells were cultured as Chapter 3 described. To make nuclear extract, Kc167 cells were washed three times in wash buffer (3.75 mM Tris pH 7.4, 0.05 mM spermine, 0.125 mM spermidine, 0.5 mM EDTA-KOH pH 7.4, 20 mM KCl, 0.5% thiodiglycol, 0.1 mM PMSF, 1 μg/mL leupeptin, 1 μg/mL chymostatin, 1 μg/mL pepstatin and 0.5 μg/mL aprotinin). Spin down the pellet at 4℃ at 2k×g for 10 min, 7 min, and 7 min, respectively. Washed cells were homogenized for 13 times in homogenization buffer (wash buffer+ 0.05%

empigen) with a Dounce homogenizer and pestle B. Nuclei were washed for three times in homogenization buffer and pelleted at 4 ℃ by centrifugation of 2k ×g, 1.5k ×g, and 1k ×g, respectively. Nuclei were resuspended in 0.5ml nuclear extraction buffer NEB. 20 (10 mM HEPES pH 7.6, 20 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, 0.2 mM PMSF, 1 μg/mL leupeptin, 1 μg/mL chymostatin, 1 μg/mL pepstatin and 0.5 μg/mL aprotinin) and mixed properly. Add NEB. 20 till total volume to 2 mL. Then add the equal volume of NEB. 700 (10 mM HEPES pH 7.6, 700 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, 0.2 mM PMSF, 1 μg/mL leupeptin, 1 μg/mL chymostatin, 1 μg/mL pepstatin and 0.5 μg/mL aprotinin) and mix immediately. Agitate gently on ice for 30 min. Extracts were ultracentrifuged at 40k ×g for 1h at 4 ℃. The supernatant, as nuclear extract, were aliquoted, flash frozen and stored at -80 ℃. The same procedure was used for *Drosophila* S2 cell nuclear extract preparation.

4.2.2 Electrophoretic Mobility Shift Assay (EMSA)

DNA fragments used in EMSA are from pSyn series plasmids *Bam*HI/*Bgl*II double digestion and gel purified with Zymoclean Gel DNA Recovery Kit (Zymo Research).The DNA fragments were end labeled with [γ-^{32}P] ATP by Klenow. End-labeled DNA fragments were incubated with S2 and Kc cell nuclear extract for 10 min at room temperature. The salt concentration was controlled between 100 mM to 150 mM in the binding reaction. In order to lower nuclear extracts or purified protein salt concentration to 100 mM, dialysis may be performed as the manufacture described (Biotech Membranes-Spectrum Laboratories, Inc). The amount of proteins used in EMSA was determined by giving roughly 50% DNA fragment shift. These reaction mixtures were loaded in a 4% polyacrylamide gels and running in 0.25 × Tris-borate-EDTA buffer for 80 min at 80 V at 4 ℃.

4.2.3 Fluorescent Electrophoresis Mobility Shift Assay (EMSA)

Probes for fluorescent EMSA were PCR amplified from pSyn-M plasmid. Two fluorescent labeled primers fp-scs'-M-FAM-5'and two normal primers fp-scs'-M-5' and fp-scs'-M-3' were used to amplify M fragment. The PCR was performed as the following procedure: 98 ℃ for 1 min; 98 ℃ for 5 sec, 57.4 ℃ for 15 sec, 72 ℃ for 15 sec, repeat this cycle for 30 times; 72 ℃ for 10 min, 10 ℃ overnight if necessary. The protocol for fluorescent EMSA is the same as radiolabeled EMSA described above except using fluorescent labeled probe instead of ^{32}P labeled probe. Different amounts of FAM labeled M fragments (0.01 pmole, 0.13 pmole, 0.47 pmole, 0.94 pmole, 1.88 pmole and 3.75 pmole) were tested in EMSA. Gels were taken pictures with Typhoon 8600 provided by LSU genomics facility center.

4.2.4 Fluorescent Footprinting

1.88 pmole (140 ng) FAM-M probes were digested by different amounts of DNase I (20 ng, 40 ng, 80 ng, 160 ng) at room temperature in digesting buffer (20 mM Tris pH 7.4, 20 mM KCl, 70 mM NaCl, 10 mM $MgCl_2$). After 100 sec, stop solution (20 mM EDTA, 2 μg/μL salmon sperm DNA) was added to block the reaction. Then ethanol precipitation was performed, following by fragment analysis. Fragment analysis was performed with 3130Xl Genetic Analyzer (Life Technologies) in LSU genomic facility center.

4.2.5 Preparation of Oligonucleotide Multimers and CNBr-activated Sepharose

DNA chromatography is an effective tool for purifying sequence specific proteins based on DNA: protein interactions (Kadonaga and Tjian, 1986). Two complementary single stranded oligonucleotides with one end phosphorylated were annealed to make double stranded oligonucleotides. These were then ligated to make multimerized oligonucleotides, followed by coupling to commercially available CNBr-activated Sepharose (GE Health Life Sciences). The

protein samples are applied to the affinity resin. Purified proteins are washed into fractions with increasing salt concentration (Kerrigan and Kadonaga, 2001).

Two 77-bp LS4D11 oligonucleotides- LS4D11-5' and LS4D11-3' with 5' ends phosphorylated- are generated for DNA chromatography. In order to make double stranded oligonucleotides, tube containing two purified oligonucleotides with one end phosphorylated was put in 95 ℃ water till the water cool down to room temperature. Ethanol precipitate double stranded oligonucleotides and dissolve the pellet in H_2O. A 99-bp LS4D21 double stranded oligonucleotides were generated by mixing 4D21-*BglSpe*-5' primer and 4D21-*BglSpe*-3' primer in boiled water. A 66-bp LS4Dmin double stranded oligonucleotides were generated by mixing LS4-D-min-*BamBgl*-5' primer and LS4-D-min-*BglBam*-3' in boiled water. Cloning three double stranded oligos LS4Dmin, LS4D11 and LS4D21 into pSyn plasmid cut with *Bam*HI/*Bgl*II. New pSyn LS4D series monomer plasmids were cut with *Sca*I/*Bam*HI and *Sca*I/*Bgl*II, respectively. Two of the fragments with *Sca*I/*Bam*HI end and *Sca*I/*Bgl*II end were ligated to generate pSyn series LS4D derivatives dimer plasmids.

The double stranded LS4D11oligo was ligated in a 100 µL reaction volume with T4-DNA ligase (New England Biolabs, Inc.) in linker buffer (66 mM Tris pH 7.6, 10 mM $MgCl_2$, 11 mM DTT, 1 mM spermidine, 4 mM ATP) for 3h at room temperature and overnight at 4 ℃. The oligonucleotide multimers were purified by two rounds of phenol/ chloroform/isoamyl alcohol, followed by ammonium acetate/ isopropanol precipitation. The final pellet was dissolved in 50 µL H_2O. Commercially available CNBr-activated Sepharose (GE Health Life Sciences) was prepared by hydrating with 500 mL of 1 mM HCl, followed by washes of 100 mL H_2O and 100 mL 10 mM KPO_4 at 4 ℃, respectively. Purified oligonucleotide multimers were coupled to the CNBr-activated sepharose resin in a solution containing 50% 10 mM KPO_4/50% resin (v/v). The

mixture was incubated at room temperature for 4.5 h by rotating. The resin coupled with DNA was centrifuged, washed with 100 mL H_2O and blocked with 100 mL of 100 mM Tris pH 8.0. Then incubate for 1.5h at room temperature. After washing at 4 ℃ with 100 mL 10 mM KPO_4 pH 8.0, 100 mL 1M KPO_4 pH 8.0, 100 mL 1M KCl, 100 mL H_2O in order, the DNA-coupled resin was stored in column storage buffer [10 mM Tris pH 8.0, 1 mM EDTA pH 8.0, 0.3 M NaCl, 0.04% (w/v) sodium azide].

4.2.6 DNA Affinity Purification

One mL of resin coupled with DNA was settled in a column for DNA affinity purification. The settled resin was equilibrated with ABN100 [25 mM HEPES pH 7.6, 0.1 mM EDTA, 0.1% NP40, 20% Glycerol, 1mM DTT, 100 mM NaCl, note: the number after ABN is the NaCl concentration in mM] twice. The salt concentration in the nuclear extracts was diluted to 100 mM, followed by centrifugation of nuclear extract at 12k×g for 10 min at 4 ℃. The supernatant of the centrifuged nuclear extracts was loaded in the column at gravity flow, followed by four rounds of 1 mL ABN100 wash. Elute the proteins from resin by adding 0.5 mL of ABN200, ABN300, ABN400, ABN500, ABN600, ABN700, ABN800 and ABN900 in a row, followed by the addition of 3 rounds of 0.5 mL ABN1000. Collect flowthrough during the wash and elution step. Add 5 mL column regeneration buffer [10 mM Tris pH 7.8, 1 mM EDTA pH 8.0, 2.5 M NaCl, 1% (v/v) NP-40] in the column twice and stir the resin to mix properly. Wash regenerated resin with 6 rounds of 5 mL column storage buffer and store in column storage buffer at 4 ℃.

4.2.7 SDS-PAGE and Western Blot

21 μL purified proteins at each fraction were loaded onto 6% polyacrylamide stacking gel and 10% polyacrylamide running gel and electrophoresis was done at room temperature at 150 V for 55 min. For western blot analysis, 12 μL of purified proteins were used. The primary affinity-

purified rabbit anti-BEAF antibody and the secondary goat anti-rabbit antibody conjugated with HRP were used at used at a 1:10000 dilution.

4.3 RESULTS

4.3.1 LS4 Region is a Potential BEAF Associated Protein Binding Site

We wanted to determine if we could detect protein binding to the LS4 sequence. To do this, we used the 225-bp M fragment which has the D site and the LS4 region. This allowed BEAF binding to serve as a positive control. Mutant probes had the LS4 sequence, the D site, or both mutated. As previously shown, mutating the D site eliminated BEAF binding. It also reduced the amount of low mobility shift, suggesting that BEAF facilitates binding of these proteins. Mutating LS4 did not affect BEAF binding, but greatly reduced the amount of low mobility (LM) complex. Perhaps the residual binding was facilitated by BEAF. Mutating both sequences

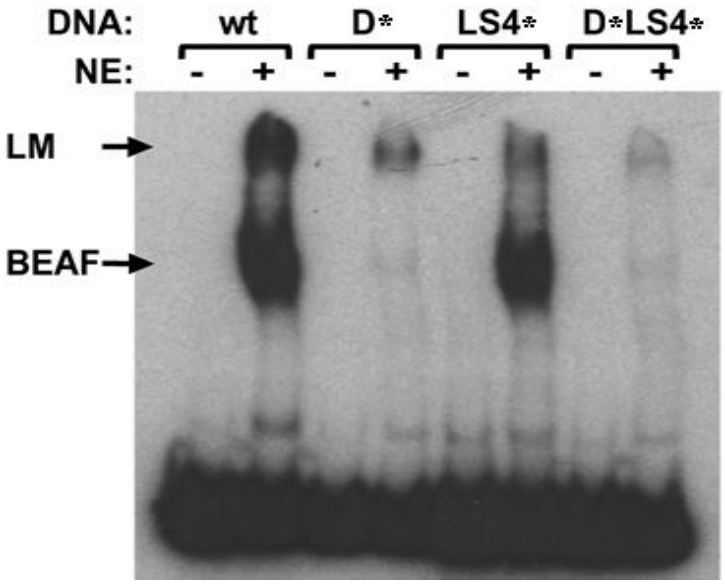

Figure 4.1 Analysis of M fragment by EMSA. NE and LM represent S2 cell nuclear extract and low mobility shift, respectively. D*, LS4* and D*LS4* represent M fragments with a D site mutation, a LS4 region mutation, or both mutations, respectively. Wt is M fragment without any mutation. Gel shift with 1.7 µg S2 cell nuclear extract (no protein in - lanes), 6 µg poly (dI-dC), and 2 fmole of indicated labeled probes (+ lanes). Reaction salt concentration was 110 mM.

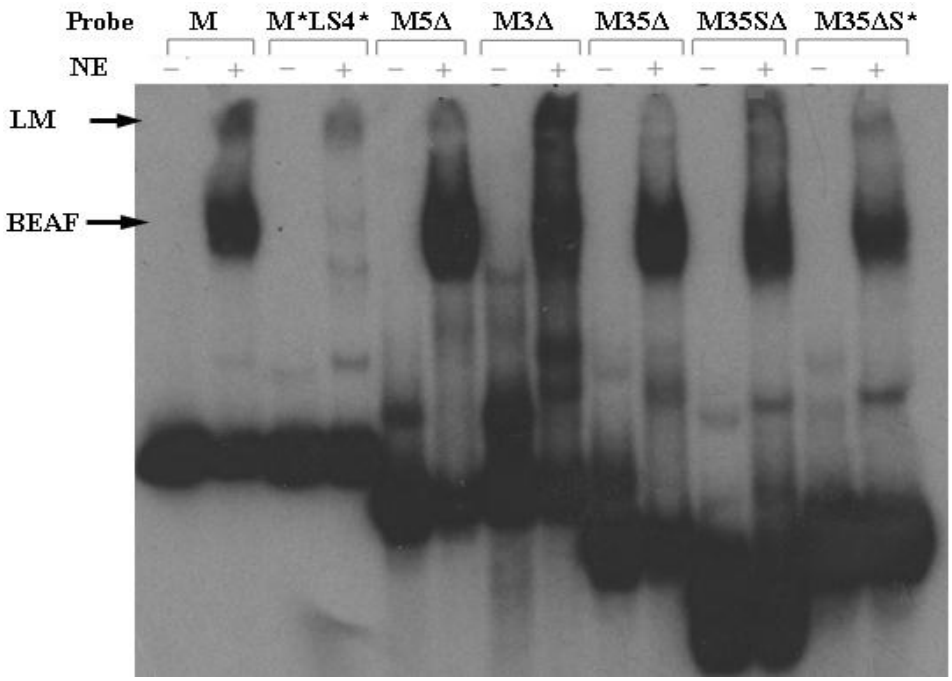

Figure 4.2 Analysis of M derivatives by EMSA. NE and LM represent S2 cell nuclear extract and low mobility shift, respectively. Star (*) and delta (Δ) represent mutation and deletion, respectively. More details of M derivatives are described in Chapter 2. Gel shift with 1.7 μg S2 cell nuclear extract (no protein in - lanes), 6 μg poly (dI-dC), and 2 fmole of indicated labeled probes (+ lanes). Reaction salt concentration was 110 mM.

essentially eliminated all binding (Figure 4.1). More EMSAs of smaller M fragments showed that deletion of only 3' end or 5' end or spacer region has subtle effect on LM shift and no effect on BEAF binding. Particularly, M35SΔ fragment without 3' end, 5' end and spacer region also recruits BEAF binding and LM binding (Figure 4.2). Dimers of each M derivatives were tested in EMSA. It was proved that dimer works better than monomer in recruiting LM binding and BEAF binding (data not shown). To summarize, there are proteins binding to LS4 sequences, and BEAF binding could help LS4 binding activity.

4.3.2 Fluorescent EMSA

Radioactivity used in EMSAs has a short half life cycle. In order to avoid using radioactivity, we explored a non-radioactive EMSA: fluorescent EMSA (Karr, 2010; Oyamada et al., 2007; Zhong and Krangel, 1999). In fluorescent EMSA, the probe is PCR amplified with fluorescent

end labeled primers. After running through a native gel, instead of drying the gel and placing the dried gel on film, the wet gel is put in a Typhoon 8600 to detect fluorescence. The M fragment labeled with FAM was used in EMSAs. In order to see free DNA, at least 0.47 pmole (35 ng) of FAM-M probe was required. However, BEAF binding shifts were not seen on the gel, probably because more BEAF was needed (Figure 4.3). In radioactive EMSA, only 2 fmole ^{32}P labeled probes are required for a decent shift. It appears that radioactive EMSA is 235 times more sensitive than fluorescent EMSA. 0.72 pmole (54 ng) FAM-M probe was tested with different amounts of bacterially expressed BEAF. Even 8 μL of bacterially expressed BEAF did not saturate BEAF binding (Figure 4.4). Evidence suggests that BEAF binds to D site in form of trimer (Gilbert et al., 2006). Then at least 2.16 pmole (69 ng) of BEAF would be needed to saturate 0.72 pmole M fragment. Therefore fluorescent EMSA does not suit in my research because of large protein amounts needed.

4.3.3 Fluorescent Footprinting

We also explored fluorescent footprinting. The fluorescent end labeled probe was digested with different amounts of DNase I at room temperature for 100 sec, followed by ethanol precipitation. Fragment analysis was performed on digested samples using the 3130Xl Genetic Analyzer. Regions without protein binding are cut by DNase I and result in peaks in this region. If bound by proteins, that region can not be cut by DNase I. Therefore, no peaks at this region are shown in the fragment analysis map. Through comparing two fragment analysis maps with and without proteins, binding sites could be identified. In the DNase I cutting optimization experiment, for 1.88 pmole (140 ng) FAM-M probe, nice peaks were shown in both 20 ng and 40 ng DNase I digestion samples; 80 ng and 160 ng DNase I digested the DNA too much (Figure 4.5). In order to saturate 1.88 pmole M fragments, at least 5.64 pmole BEAF are required. It seems too much

for purified proteins. Therefore, fluorescent footprinting is not suitable for DNA affinity purified proteins.

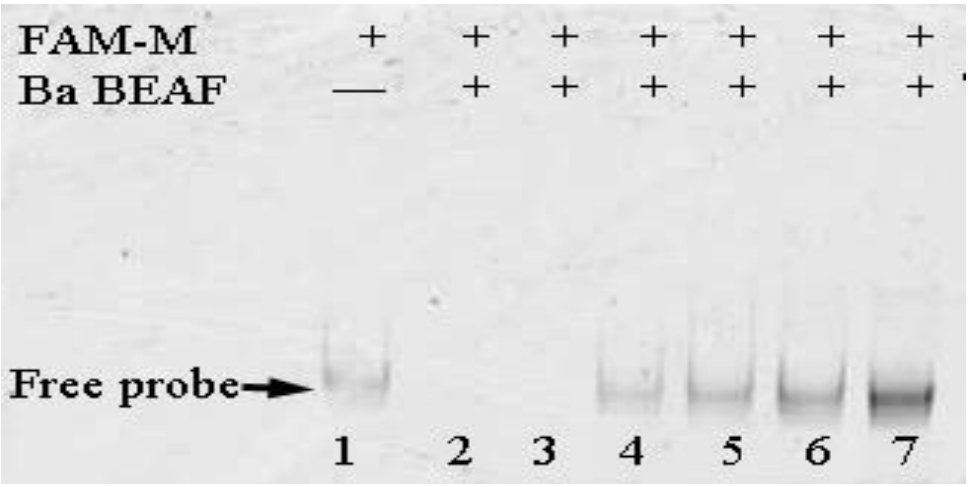

Figure 4.3 Fluorescent EMSA with different amounts of FAM-labeled M fragment. Ba BEAF is bacterially expressed BEAF. 1 µL Ba BEAF (no protein in lane 1), 6 µg poly (dI-dC) (lane 2-7), and 0.01 pmole was added. 0.13 pmole, 0.47 pmole, 0.94 pmole, 1.88 pmole and 3.75 pmole FAM-M probes were added to lane 2 to lane 7, respectively. 0.94 pmole FAM- M fragments were added in lane 1 for control. The salt concentration of binding reaction was between 100nM to 150 nM.

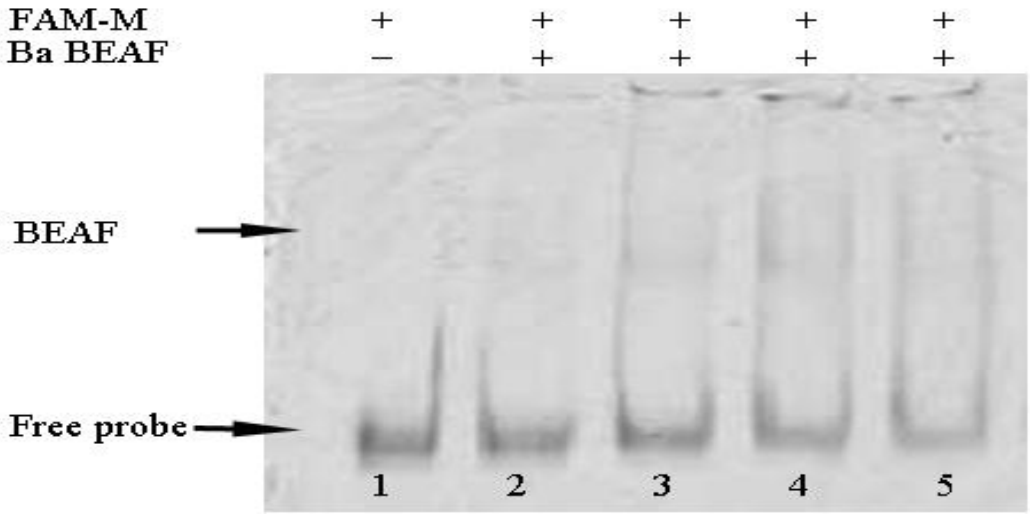

Figure 4.4 Fluorescent EMSA with different amounts of BEAF. Ba BEAF is bacterially expressed BEAF. 0.72 pmole FAM-M probes were used in each lane. 2 µL, 4 µL, 6 µL, 8 µL bacterially expressed BEAF was added to lane 2 to lane 5, respectively. The salt concentration of binding reaction was between 100nM to 150 nM.

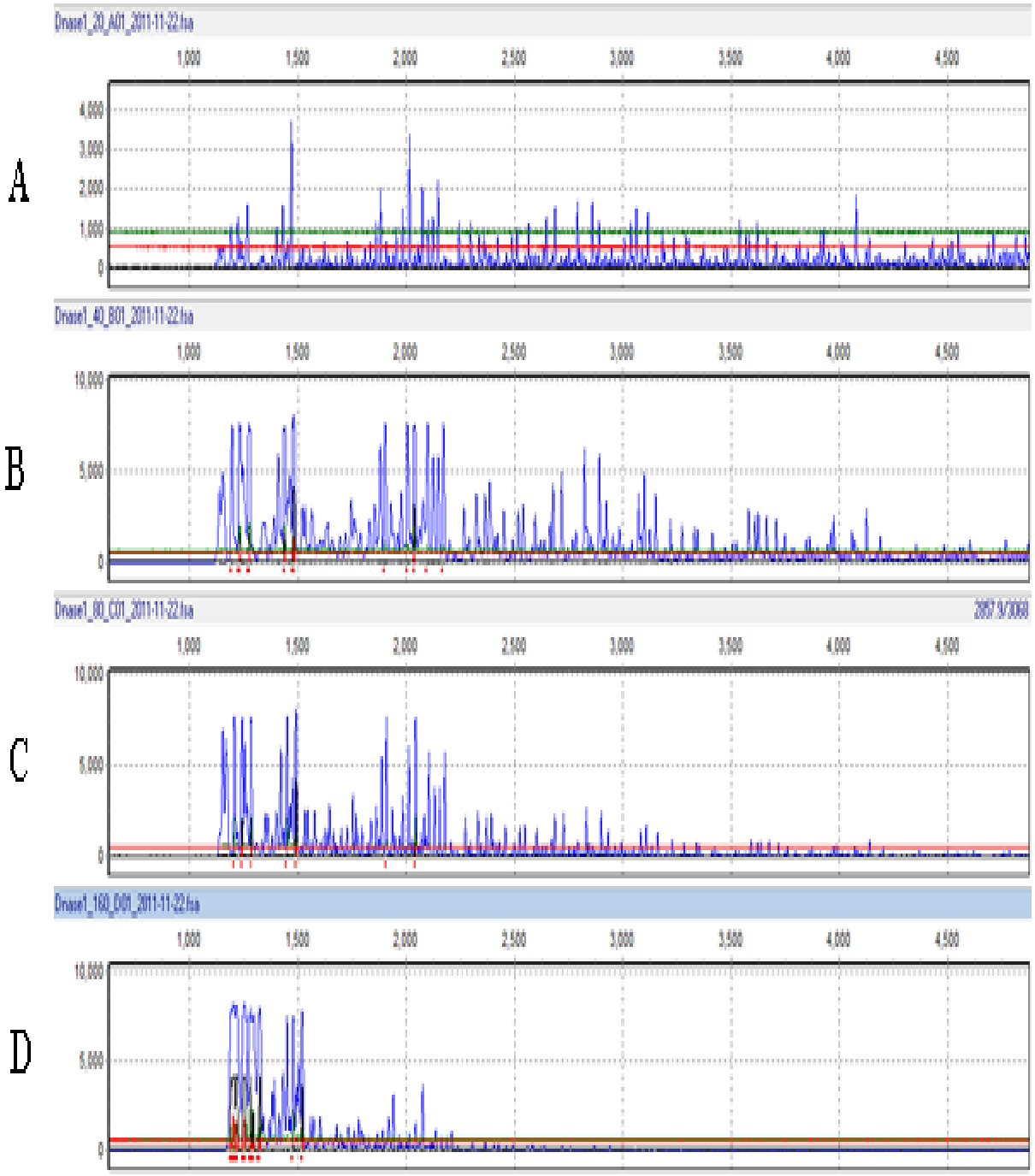

Figure 4.5 Fragment analysis of FAM-M fragment after DNase I digestion. A, B, C and D are the map of FAM-M fragment digested by 20 ng, 40 ng, 80 ng and 260 ng DNase I, respectively. 1.88 pmole FAM-M probes were used for each sample.

4.3.4 Several LS4 Binding Proteins are Purified by DNA Affinity Chromatography

Based on the above EMSA results and fly chromosomal position effect assays, proteins that function with BEAF are likely to bind to the LS4 sequence. In addition, BEAF binding assisted LM binding directly or indirectly. Based on this, we designed an oligonucleotide for DNA affinity chromatography that has both the LS4 sequence and the D site. To minimize the length of the oligo, experiments were done with oligos with spacer shorter than the 52 bp separating LS4 region and D site. The dimers of LS4Dmin, LS4D11, M35SΔ and LS4D21 have 3 bp, 11 bp, 10 bp and 21 bp of spacing between LS4 region and D site, respectively (Figure 4.6). EMSA showed that the LS4D11 dimer worked better than the LS4Dmin dimer in recruiting the LM shift. Since more BEAF binding and LM binding were detected in Kc cells than in S2 cells in EMSA, Kc cell nuclear extracts were used for DNA affinity chromatography (Figure 4.7 A and B). With an 11 bp spacer between the LS4 and D sequences, the resulting oligo was 77 bp. This LS4D11 oligo was concatenated (Figure 4.8) and coupled to commercially available CNBr-activated Sepharose for DNA affinity chromatography.

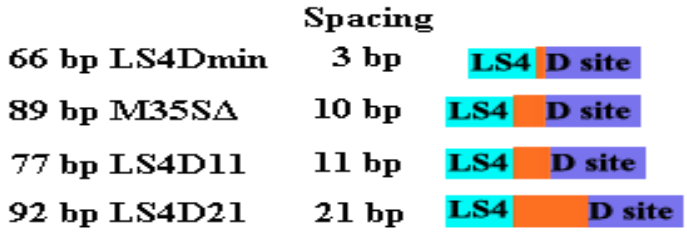

Figure 4.6 Diagram of oligonucleotide candidates for DNA affinity purification. The orange box represents the spacing between the D site and LS4 region. 5' end and 3' end are not shown.

Nuclear proteins from Kc cells were bound to the LS4D11 resin and eluted using 100 mM steps of NaCl. Fractions were subjected to SDS-PAGE and visualized by Coomassie staining or Western blot analysis with anti-BEAF antibodies. Proteins that stuck to the column mainly eluted at 300-600 mM NaCl with a peak at 400 mM (Figure 4.9), while BEAF mainly eluted at 400-600

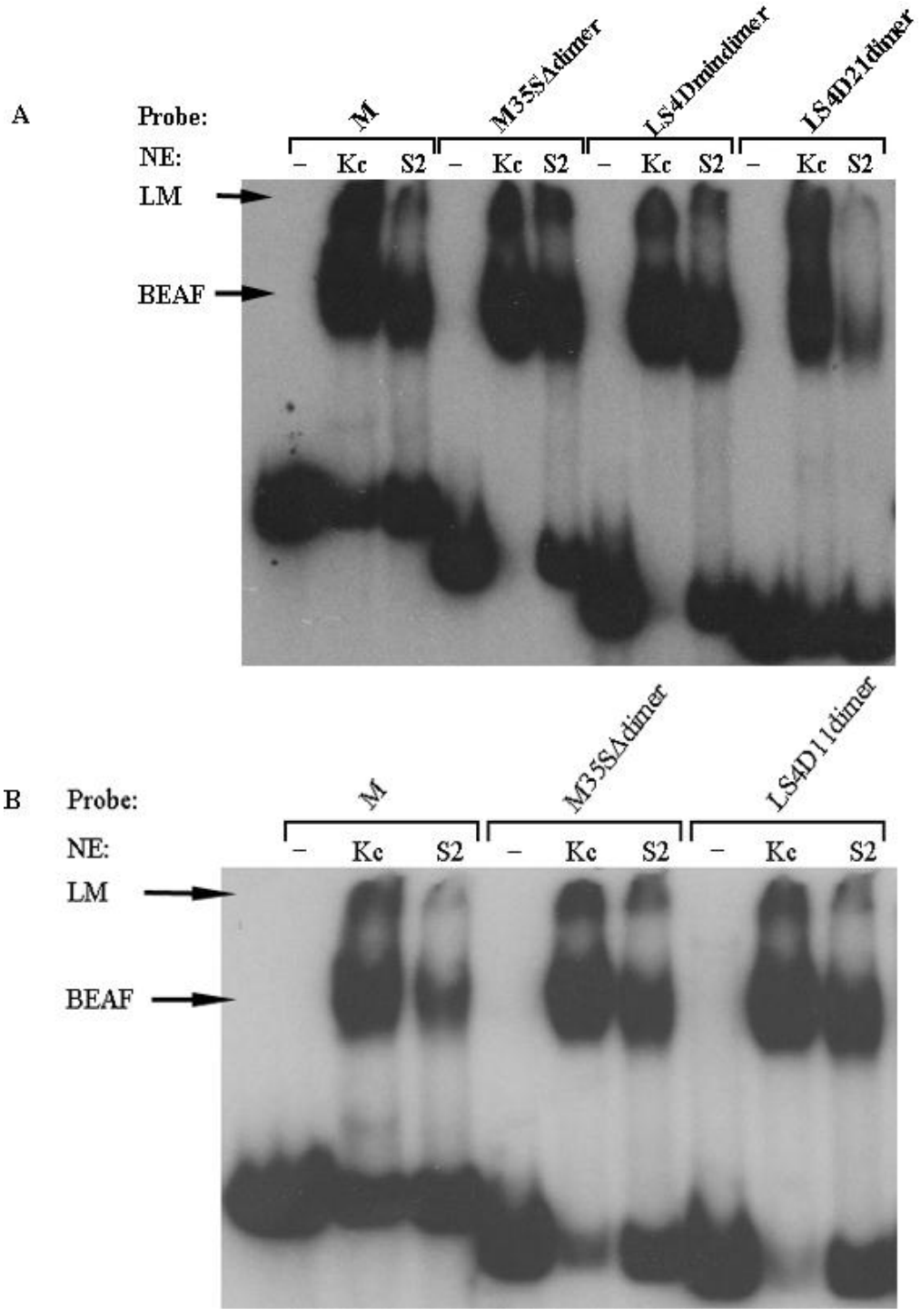

Figure 4.7 Analysis of LS4D derivatives binding activity by EMSA. 1.8 μg Kc cell nuclear extracts and S2 cell nuclear extracts were added in Kc and S2 lane, respectively. 6 μg poly (dI-dC), and 2 fmole of indicated labeled probes were added in Kc and S2 lanes. 126-bp LS4Dmin dimer, 153-bp LS4D11 dimer, 186-bp LS4D21 dimer, 172-bp M35SΔ dimer have 3-bp, 11-bp, 21-bp and 10-bp spacers between LS4 region and D site, respectively. Reaction salt concentration was 110 mM.

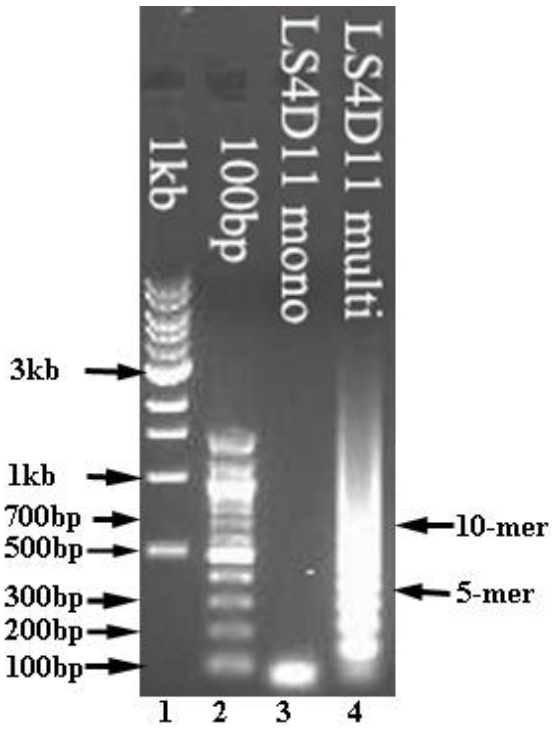

Figure 4.8 Analysis of the ligated LS4D11 oligo by 1.2% agarose gel electrophoresis. The DNA was visualized by ethidium bromide staining and ultraviolet light fluorescence. Lane 1 and lane 2 are DNA markers. Lane 3 is unligated 77-bp LS4D11 oligonucleotides (250 ng). Lane 4 is ligated LS4D11 oligonucleotides (0.5 μL). The estimated migrateion of multimers is indicated. TAE buffer was used.

mM NaCl with a peak at 500 mM (Figure 4.10). It was difficult to detect the BEAF shift and LM shift after dialyzing the proteins to reduce the salt concentration. Instead, smears were detected (Figure 4.11). To determine if proteins were lost during dialysis, SDS-PAGE of dialyzed proteins was done. The amount of proteins did not decrease, indicating it was DNA binding activity loss that caused the failed EMSA after dialysis (Figure 4.12). Next, we combined the dialyzed ABN400, ABN500 and ABN600 fractions together, and ran this through the affinity column a second time. The results of the second purification were consistent with the first purification. The multiple bands in the ABN 300 and ABN400 lanes were the LS4 binding protein candidates (Figure 4.13). These results have been repeated with several nuclear extracts. EMSA and

footprinting experiments need to be done to determine specificity for the LS4 sequence. If specific, the proteins can be identified by mass spectrometry.

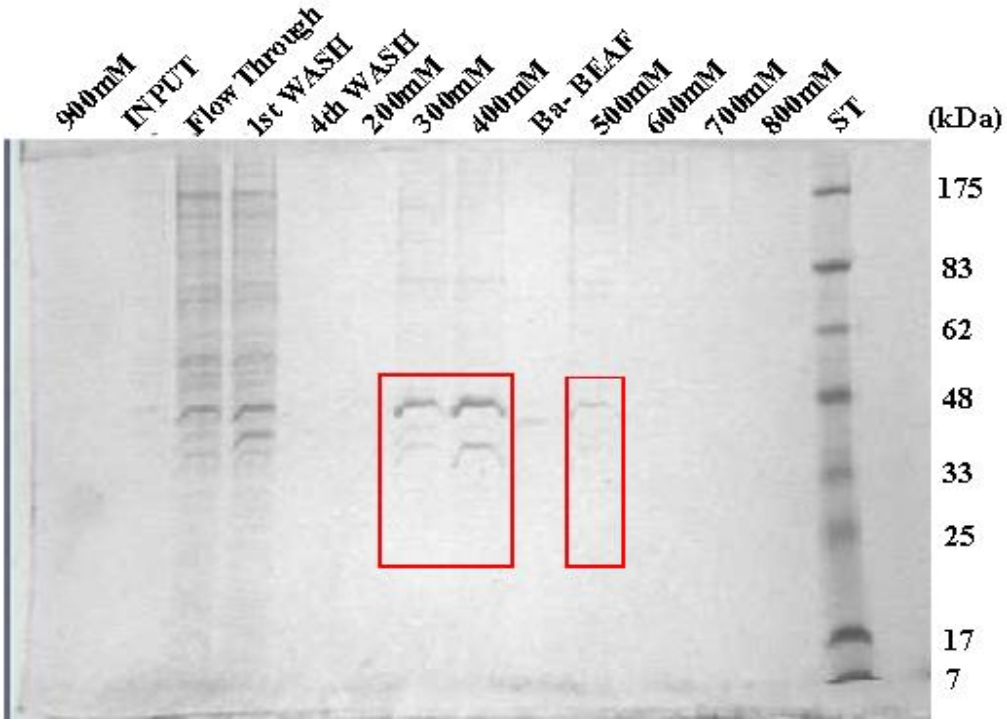

Figure 4.9 Coomassie staining of 1st purification of LS4 binding proteins from Kc cell nuclear extracts. 21 µL purified proteins were loaded, except that Only 2 µL input, 12 µL bacterially expressed BEAF (Ba BEAF) and 9 µL protein standards were loaded. Bands in red boxes were proteins which may bind to LS4 region.

4.4 DISCUSSION

EMSA has demonstrated that LS4 is bound by proteins and LS4 binding activity is affected by BEAF binding to the D site. However, it is still unclear if LS4 binding proteins are LS4 specific or how BEAF affects LS4 binding activity. Footprinting could be done to answer the first question. We explored using fluorescent detection of DNA binding. DNase I-generated fragments were detectable when using 1.88 pmole of FAM-M DNA. This would require 5.64 pmole (180 ng) of BEAF to saturate binding. In contrast, radioactive detection would require only 0.2 ng of BEAF. The same is true for EMSAs, where 0.47 pmole of probe is needed for

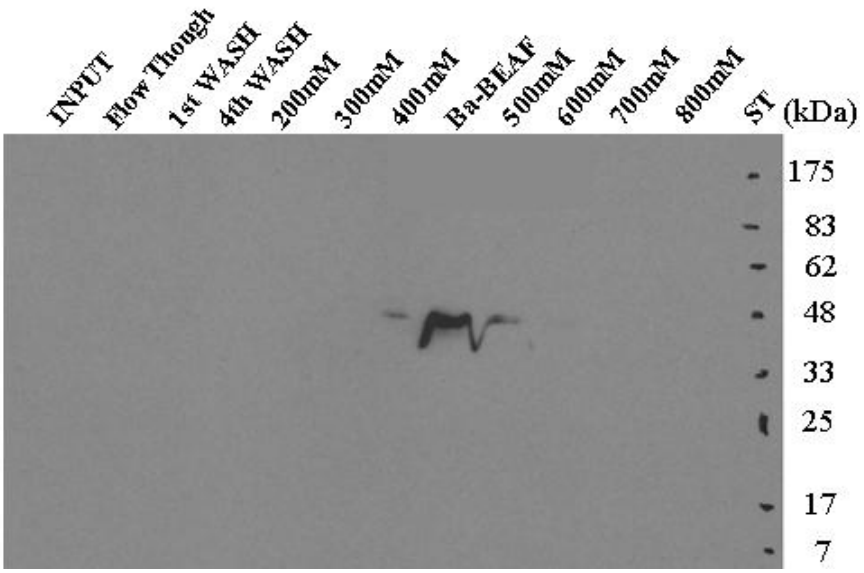

Figure 4.10 Western blot of DNA affinity purified protein against BEAF. 4 µL input, 12 µL bacterially expressed BEAF (Ba-BEAF) and 9 µL protein standard were loaded. 12 µL purified proteins were loaded.

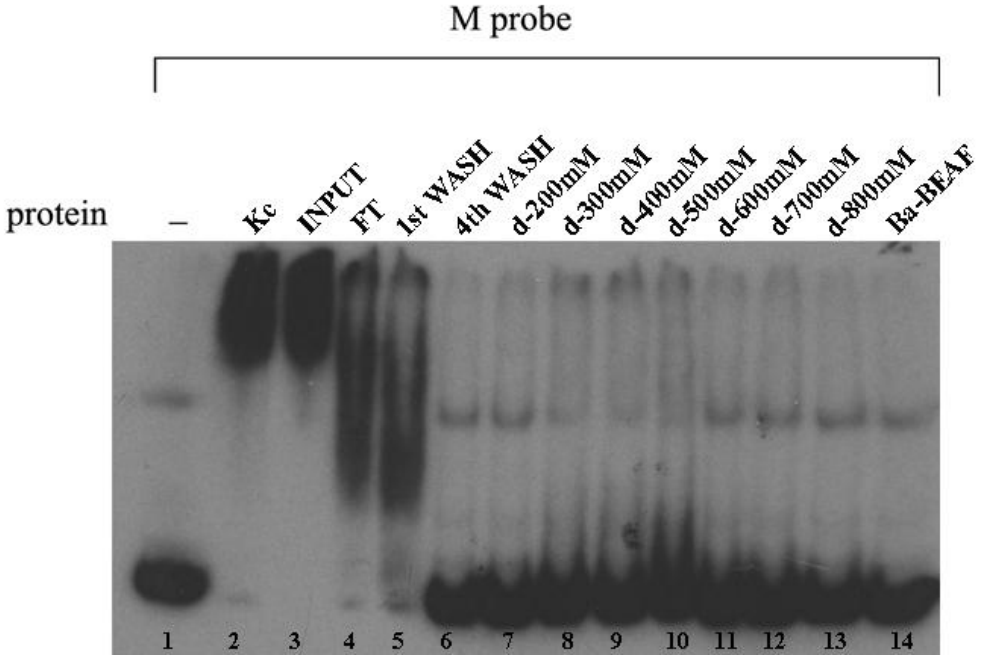

Figure 4.11 EMSA of dialyzed 1st purification proteins with M probe. 3 µL dialyzed affinity purified proteins were loaded to lane 4 to lane 13, respectively. 1 µL bacterially expressed BEAF (Ba-BEAF) was loaded in lane 14. 1µg poly (dI-dC) was added in lane 2 to lane 14. The concentration of binding reaction was 100 mM to 150 mM.

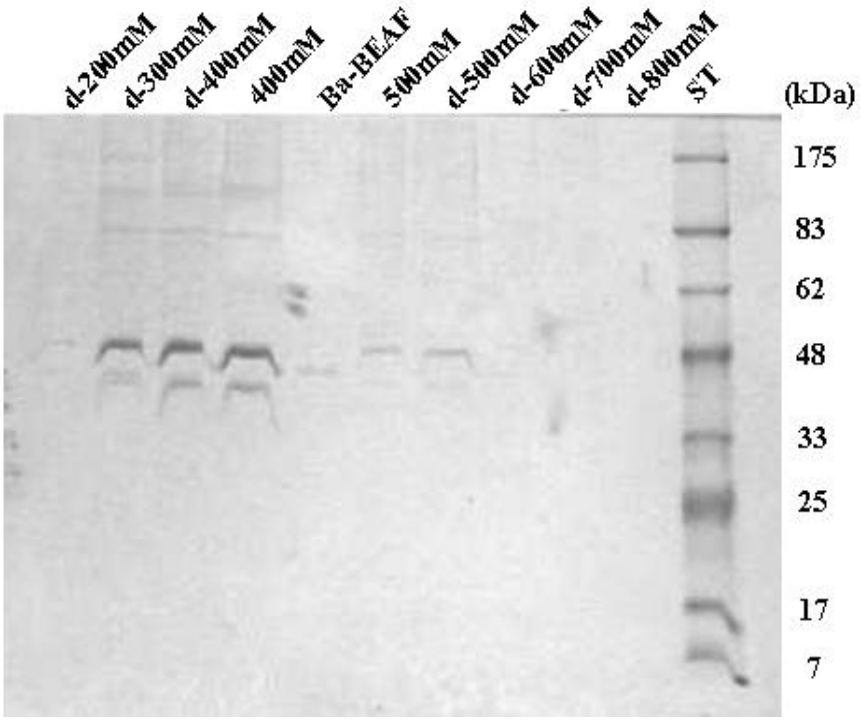

Figure 4.12 Coomassie staining of dialyzed 1st purification of LS4 binding proteins from Kc cell nuclear extracts. The same amount of probe, proteins and poly (dI-dC) as described in Figure 4.7. The salt concentration of binding reaction is 100 mM to 150 mM.

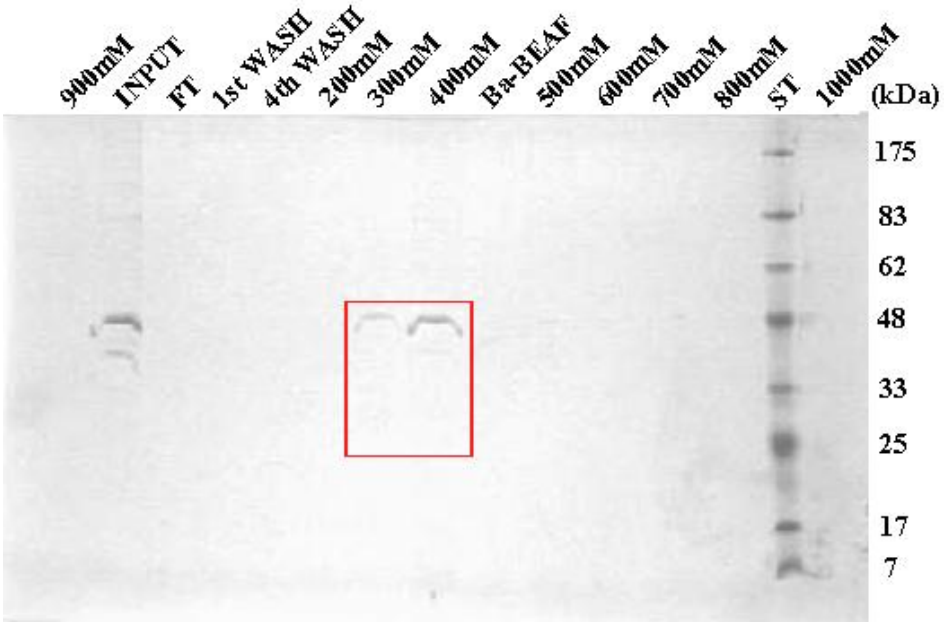

Figure 4.13 Coomassie staining of 2nd purification of of LS4 binding proteins from Kc cell nuclear extracts. The same amount of probe, proteins and poly (dI-dC) as described in Figure 4.7. The salt concentration of binding reaction is 100 mM to 150 mM. Bands in red boxes were consistent with 1st purification bands at 300 and 400 mM salt concentration.

fluorescent detection compared to 2 fmole probe with radioactivity. Fluorescent detection might be suitable for bacterially expressed proteins, but not for the small amounts of proteins in nuclear extracts.

In the DNA affinity purification, it was exciting that a few bands were seen consistently for different nuclear extract preps after running through the affinity column, indicating those bands may be LS4 binding proteins. Surprisingly, BEAF was washed off at ABN400, ANB500 and ABN600, mainly at ABN500. However, Coomassie staining revealed multiple bands are at ABN300, ABN400 and ABN500, mainly at ABN400. It is not clear if these bands are background proteins rather than LS4 binding proteins. In order to reduce the unspecific binding to resin, non- specific competitor dIdC should be added in the sample before running the affinity column.

4.5 FUTURE WORK

EMSA of affinity purified proteins needs to be done to determine how much purified protein is needed to saturate binding to the probe. Footprinting could then be done to see if LS4 sequences are protected. The non-specific competitor poly (dI-dC) should be added to the nuclear extract sample before DNA affinity chromatography to see if the proteins are still purified. If footprinting demonstrates LS4 is bound by some proteins and proteins washed off from affinity resin are not non-specific binding proteins; mass spec could be done to identify the proteins.

CHAPTER 5
SUMMARY

Previous study has demonstrated that BEAF was essential for scs' insulator activity and clustered CGATA motifs are necessary for BEAF binding. The D site in scs' is a high affinity BEAF binding site that plays a pivotal role in the insulator function of scs' and the M fragment. Mutation of the D site eliminates insulator function in position effect assays. Position effect assays of more M derivatives revealed that the LS4 region mutation compromises M2 insulator activity, indicating the LS4 region is essential for full insulator function. Based on these results, LS4 sequences help BEAF work better in insulator activity. We designed injection plasmids of M derivatives containing the D site and LS4 region to test and determine the minimal scs' sequences for proper insulator activity. To simplify the assay, the *Φ*C31 mediated site specific integration system will be used for fly injections. Since it uses site specific integration rather than the random integration of P elements, only two fly lines are enough for the planned experiments. Because *attP* fly lines available from Bloomington Drosophila Stock Center did not show strong position effects, we have been trying to make our own *attP* fly lines. Currently, 3 out of 15 transgenic fly lines show decent position effects. However, they are either homozygous lethal or sterile. I am still screening transgenic flies to find lines with strong position effects.

In contrast to other insulator binding proteins like Su(Hw), dCTCF and CP190, more than 85% of BEAF binding sites in the *Drosophila* genome were within 300 bp of transcription start sites, indicating that BEAF may play a role in promoter activity in addition to insulator activity. An example is scs' which encompassed two divergent genes, *aurora* and *CG3281*. Moreover, BEAF knock-out decreased BEAF associated gene expression levels. In order to see if M, which has insulator activity, has promoter activity and how the insulator activity relates with promoter activity, a promoter activity assay in transfected cultured cells is being performed. This assay

tests whether the M fragment and mutant versions can function as promoters for transgenes encoding GFP and RFP tagged nucleolar protein. It was observed that M has promoter activity. In the future, we will test other M derivatives for promoter activity in S2 cells. If we are able to identify minimal scs' insulator sequences in transgenic flies, RT-PCR could be done to see if the insulator also has promoter activity, and to see if these results agree with the tissue culture assay.

EMSA of M derivatives have shown that proteins bind to the LS4 region, and that BEAF binding assists LS4 binding. This is consistent with the fact that LS4 strengthens insulator activity. Fluorescent EMSA and footprinting are much less sensitive than radiolabeled EMSA and footprinting. In particular, fluorescent EMSA and footprinting are not good options for low amounts of proteins purified from nuclear extracts. DNA affinity chromatography was performed to purify LS4 binding proteins. Bands of interest were observed after two rounds of affinity purification. However, footprinting and mass spectrometry need to be done to further identify if they are LS4 specific binding proteins and what proteins they are.

REFERENCES

Adryan, B., Woerfel, G., Birch-Machin, I., Gao, S., Quick, M., Meadows, L., Russell, S., and White, R. (2007). Genomic mapping of Suppressor of Hairy-wing binding sites in Drosophila. Genome biology *8*, R167.

Aravind, L. (2000). The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. Trends in biochemical sciences *25*, 421-423.

Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Muller, M., Shanower, G., Schedl, P., Gyurkovics, H., and Karch, F. (2000). The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. Development *127*, 779-790.

Bartkuhn, M., Straub, T., Herold, M., Herrmann, M., Rathke, C., Saumweber, H., Gilfillan, G.D., Becker, P.B., and Renkawitz, R. (2009). Active promoters and insulators are marked by the centrosomal protein 190. The EMBO journal *28*, 877-888.

Baxley, R.M., Soshnev, A.A., Koryakov, D.E., Zhimulev, I.F., and Geyer, P.K. (2011). The role of the Suppressor of Hairy-wing insulator protein in Drosophila oogenesis. Developmental biology *356*, 398-410.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Belozerov, V.E., Majumder, P., Shen, P., and Cai, H.N. (2003). A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of Drosophila. The EMBO journal *22*, 3113-3121.

Biggin, M.D., and Tjian, R. (1988). Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts. Cell *53*, 699-711.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proceedings of the National Academy of Sciences of the United States of America *104*, 3312-3317.

Blanton, J., Gaszner, M., and Schedl, P. (2003). Protein:protein interactions and the pairing of boundary elements in vivo. Genes & development *17*, 664-675.

Brasset, E., Hermant, C., Jensen, S., and Vaury, C. (2010). The Idefix enhancer-blocking insulator also harbors barrier activity. Gene *450*, 25-31.

Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., and Dorn, R. (2000). Genetic and molecular complexity of the position effect variegation modifier mod(mdg4) in Drosophila. Genetics *155*, 141-157.

Bulger, M., and Groudine, M. (2010). Enhancers: the abundance and function of regulatory sequences beyond promoters. Developmental biology *339*, 250-257.

Bushey, A.M., Ramos, E., and Corces, V.G. (2009). Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. Genes & development *23*, 1338-1350.

Butcher, R.D., Chodagam, S., Basto, R., Wakefield, J.G., Henderson, D.S., Raff, J.W., and Whitfield, W.G. (2004). The Drosophila centrosome-associated protein CP190 is essential for viability but not for cell division. Journal of cell science *117*, 1191-1199.

Capelson, M., and Corces, V.G. (2005). The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. Molecular cell *20*, 105-116.

Chodagam, S., Royou, A., Whitfield, W., Karess, R., and Raff, J.W. (2005). The centrosomal protein CP190 regulates myosin function during early Drosophila development. Current biology : CB *15*, 1308-1313.

Comet, I., Schuettengruber, B., Sexton, T., and Cavalli, G. (2011). A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. Proceedings of the National Academy of Sciences of the United States of America *108*, 2294-2299.

Cuvier, O., Hart, C.M., and Laemmli, U.K. (1998). Identification of a class of chromatin boundary elements. Molecular and cellular biology *18*, 7478-7486.

Dickson, J., Gowher, H., Strogantsev, R., Gaszner, M., Hair, A., Felsenfeld, G., and West, A.G. (2010). VEZF1 elements mediate protection from DNA methylation. PLoS genetics *6*, e1000804.

Donze, D., and Kamakaka, R.T. (2001). RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. The EMBO journal *20*, 520-531.

Dorsett, D. (1993). Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of Drosophila. Genetics *134*, 1135-1144.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. Nature *371*, 806-808.

Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Molecular and cellular biology *22*, 3820-3831.

Faucheux, M., Roignant, J.Y., Netter, S., Charollais, J., Antoniewski, C., and Theodore, L. (2003). batman Interacts with polycomb and trithorax group genes and encodes a BTB/POZ protein that is included in a complex containing GAGA factor. Molecular and cellular biology *23*, 1181-1195.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nature reviews Genetics *7*, 703-713.

Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. Genes & development *13*, 2098-2107.

Gause, M., Morcillo, P., and Dorsett, D. (2001). Insulation of enhancer-promoter communication by a gypsy transposon insert in the Drosophila cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. Molecular and cellular biology *21*, 4807-4817.

Gerasimova, T.I., and Corces, V.G. (1998). Polycomb and trithorax group proteins mediate the function of a chromatin insulator. Cell *92*, 511-521.

Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V., Simonova, O., and Corces, V.G. (1995). A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. Cell *82*, 587-597.

Gerasimova, T.I., Lei, E.P., Bushey, A.M., and Corces, V.G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. Molecular cell *28*, 761-772.

Geyer, P.K., and Corces, V.G. (1992). DNA position-specific repression of transcription by a Drosophila zinc finger protein. Genes & development *6*, 1865-1873.

Gilbert, M.K., Tan, Y.Y., and Hart, C.M. (2006). The Drosophila boundary element-associated factors BEAF-32A and BEAF-32B affect chromatin structure. Genetics *173*, 1365-1375.

Golovnin, A., Melnikova, L., Volkov, I., Kostuchenko, M., Galkin, A.V., and Georgiev, P. (2008). 'Insulator bodies' are aggregates of proteins but not of insulators. EMBO reports *9*, 440-445.

Gyurkovics, H., Gausz, J., Kummer, J., and Karch, F. (1990). A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation. The EMBO journal *9*, 2579-2585.

Hart, C.M., Cuvier, O., and Laemmli, U.K. (1999). Evidence for an antagonistic relationship between the boundary element-associated factor BEAF and the transcription factor DREF. Chromosoma *108*, 375-383.

Hart, C.M., Zhao, K., and Laemmli, U.K. (1997). The scs' boundary element: characterization of boundary element-associated factors. Molecular and cellular biology *17*, 999-1009.

Heath, H., Ribeiro de Almeida, C., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.W., Gribnau, J., Renkawitz, R., Grosveld, F., *et al.* (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. The EMBO journal *27*, 2839-2850.

Holdridge, C., and Dorsett, D. (1991). Repression of hsp70 heat shock gene transcription by the suppressor of hairy-wing protein of Drosophila melanogaster. Molecular and cellular biology *11*, 1894-1900.

Horn, C., and Handler, A.M. (2005). Site-specific genomic targeting in Drosophila. Proceedings of the National Academy of Sciences of the United States of America *102*, 12483-12488.

Jiang, N., Emberly, E., Cuvier, O., and Hart, C.M. (2009). Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in Drosophila melanogaster links BEAF to transcription. Molecular and cellular biology *29*, 3556-3568.

Jimenez, M., and Goday, C. (1993). A centrosome-associated antibody from Drosophila melanogaster reveals a new microtubule-dependent structure in the equatorial zone of Parascaris univalens embryos. Journal of cell science *106 (Pt 3)*, 719-730.

Kadonaga, J.T., and Tjian, R. (1986). Affinity purification of sequence-specific DNA binding proteins. Proceedings of the National Academy of Sciences of the United States of America *83*, 5889-5893.

Karch, F., Galloni, M., Sipos, L., Gausz, J., Gyurkovics, H., and Schedl, P. (1994). Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of Drosophila melanogaster. Nucleic acids research *22*, 3138-3146.

Karr, E.A. (2010). The methanogen-specific transcription factor MsvR regulates the fpaA-rlp-rub oxidative stress operon adjacent to msvR in Methanothermobacter thermautotrophicus. Journal of bacteriology *192*, 5914-5922.

Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell *64*, 941-950.

Kellum, R., and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Molecular and cellular biology *12*, 2424-2431.

Kerrigan, L.A., and Kadonaga, J.T. (2001). Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography. In Current protocols in protein science.

Kingston, R.E. (2003). Introduction of DNA into Mammalian Cells. In Current Protocols in Molecular Biology.

Maeda, R.K., and Karch, F. (2007). Making connections: boundaries and insulators in Drosophila. Current opinion in genetics & development *17*, 394-399.

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S.E., and Perrimon, N. (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nature genetics *40*, 476-483.

Mihaly, J., Hogga, I., Gausz, J., Gyurkovics, H., and Karch, F. (1997). In situ dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element. Development *124*, 1809-1820.

Mishra, K., Chopra, V.S., Srinivasan, A., and Mishra, R.K. (2003). Trl-GAGA directly interacts with lola like and both are part of the repressive complex of Polycomb group of genes. Mechanisms of development *120*, 681-689.

Modolell, J., Bender, W., and Meselson, M. (1983). Drosophila melanogaster mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. Proceedings of the National Academy of Sciences of the United States of America *80*, 1678-1682.

Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heinl, N., Bardenhagen, I., Leers, J., White, R.A., Renkawitz-Pohl, R., Saumweber, H., *et al.* (2007). The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. The EMBO journal *26*, 4203-4214.

Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S.T., Munhall, A., Grewe, B., Bartkuhn, M., Arnold, R., *et al.* (2005). CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO reports *6*, 165-170.

Morcillo, P., Rosen, C., Baylies, M.K., and Dorsett, D. (1997). Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in Drosophila. Genes & development *11*, 2729-2740.

Muravyova, E., Golovnin, A., Gracheva, E., Parshikov, A., Belenkaya, T., Pirrotta, V., and Georgiev, P. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. Science *291*, 495-498.

Mutskov, V.J., Farrell, C.M., Wade, P.A., Wolffe, A.P., and Felsenfeld, G. (2002). The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. Genes & development *16*, 1540-1554.

Negre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., *et al.* (2010). A comprehensive map of insulator elements for the Drosophila genome. PLoS genetics *6*, e1000814.

Noma, K., Cam, H.P., Maraia, R.J., and Grewal, S.I. (2006). A role for TFIIIC transcription factor complex in genome organization. Cell *125*, 859-872.

Oberstein, A., Pare, A., Kaplan, L., and Small, S. (2005). Site-specific transgenesis by Cre-mediated recombination in Drosophila. Nature methods *2*, 583-585.

Ohtsuki, S., and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. Genes & development *12*, 3325-3330.

Ong, C.T., and Corces, V.G. (2009). Insulators as mediators of intra- and inter-chromosomal interactions: a common evolutionary theme. Journal of biology *8*, 73.

Oyamada, T., Yokoyama, K., Morinaga, M., Suzuki, M., and Makino, K. (2007). Expression of Escherichia coli DcuS-R two-component regulatory system is regulated by the secondary internal promoter which is activated by CRP-cAMP. J Microbiol *45*, 234-240.

Pai, C.Y., Lei, E.P., Ghosh, D., and Corces, V.G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. Molecular cell *16*, 737-748.

Parnell, T.J., Kuhn, E.J., Gilmore, B.L., Helou, C., Wold, M.S., and Geyer, P.K. (2006). Identification of genomic sites that bind the Drosophila suppressor of Hairy-wing insulator protein. Molecular and cellular biology *26*, 5983-5993.

Peifer, M., and Bender, W. (1988). Sequences of the gypsy transposon of Drosophila necessary for its effects on adjacent genes. Proceedings of the National Academy of Sciences of the United States of America *85*, 9650-9654.

Petersen, N.S., Lankenau, D.H., Mitchell, H.K., Young, P., and Corces, V.G. (1994). forked proteins are components of fiber bundles present in developing bristles of Drosophila melanogaster. Genetics *136*, 173-182.

Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z., and Lahn, B.T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. PloS one *5*, e10611.

Raff, J.W., Kellogg, D.R., and Alberts, B.M. (1993). Drosophila gamma-tubulin is part of a complex containing two previously identified centrosomal MAPs. The Journal of cell biology *121*, 823-835.

Ramos, E., Torre, E.A., Bushey, A.M., Gurudatta, B.V., and Corces, V.G. (2011). DNA topoisomerase II modulates insulator function in Drosophila. PloS one *6*, e16562.

Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proceedings of the National Academy of Sciences of the United States of America *99*, 6883-6888.

Rollins, R.A., Morcillo, P., and Dorsett, D. (1999). Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics *152*, 577-593.

Roseman, R.R., Pirrotta, V., and Geyer, P.K. (1993). The su(Hw) protein insulates expression of the Drosophila melanogaster white gene from chromosomal position-effects. The EMBO journal *12*, 435-442.

Roy, S., Gilbert, M.K., and Hart, C.M. (2007). Characterization of BEAF mutations isolated by homologous recombination in Drosophila. Genetics *176*, 801-813.

Rubin, G.M., and Spradling, A.C. (1982). Genetic transformation of Drosophila with transposable element vectors. Science *218*, 348-353.

Rutledge, B.J., Mortin, M.A., Schwarz, E., Thierry-Mieg, D., and Meselson, M. (1988). Genetic interactions of modifier genes and modifiable alleles in Drosophila melanogaster. Genetics *119*, 391-397.

Schweinsberg, S., Hagstrom, K., Gohl, D., Schedl, P., Kumar, R.P., Mishra, R., and Karch, F. (2004). The enhancer-blocking activity of the Fab-7 boundary from the Drosophila bithorax complex requires GAGA-factor-binding sites. Genetics *168*, 1371-1384.

Scott, K.C., Taubman, A.D., and Geyer, P.K. (1999). Enhancer blocking by the Drosophila gypsy insulator depends upon insulator anatomy and enhancer strength. Genetics *153*, 787-798.

She, W., Lin, W., Zhu, Y., Chen, Y., Jin, W., Yang, Y., Han, N., Bian, H., Zhu, M., and Wang, J. (2010). The gypsy insulator of Drosophila melanogaster, together with its binding protein suppressor of Hairy-wing, facilitate high and precise expression of transgenes in Arabidopsis thaliana. Genetics *185*, 1141-1150.

Siegal, M.L., and Hartl, D.L. (1996). Transgene Coplacement and high efficiency site-specific recombination with the Cre/loxP system in Drosophila. Genetics *144*, 715-726.

Simms, T.A., Dugas, S.L., Gremillion, J.C., Ibos, M.E., Dandurand, M.N., Toliver, T.T., Edwards, D.J., and Donze, D. (2008). TFIIIC binding sites function as both heterochromatin barriers and chromatin insulators in Saccharomyces cerevisiae. Eukaryotic cell *7*, 2078-2086.

Smith, P.A., and Corces, V.G. (1992). The suppressor of Hairy-wing binding region is required for gypsy mutagenesis. Molecular & general genetics : MGG *233*, 65-70.

Soeller, W.C., Oh, C.E., and Kornberg, T.B. (1993). Isolation of cDNAs encoding the Drosophila GAGA transcription factor. Molecular and cellular biology *13*, 7961-7970.

Soshnev, A.A., He, B., Baxley, R.M., Jiang, N., Hart, C.M., Tan, K., and Geyer, P.K. (2012). Genome-wide studies of the multi-zinc finger Drosophila Suppressor of Hairy-wing protein in the ovary. Nucleic acids research.

Spana, C., and Corces, V.G. (1990). DNA bending is a determinant of binding specificity for a Drosophila zinc finger protein. Genes & development *4*, 1505-1515.

Spana, C., Harrison, D.A., and Corces, V.G. (1988). The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. Genes & development *2*, 1414-1423.

Spradling, A.C., and Rubin, G.M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. Science *218*, 341-347.

Tchurikov, N.A., Kretova, O.V., Moiseeva, E.D., and Sosin, D.V. (2009). Evidence for RNA synthesis in the intergenic region between enhancer and promoter and its inhibition by insulators in Drosophila melanogaster. Nucleic acids research *37*, 111-122.

Thorpe, H.M., Wilson, S.E., and Smith, M.C. (2000). Control of directionality in the site-specific recombination system of the Streptomyces phage phiC31. Molecular microbiology *38*, 232-241.

Udvardy, A., Maine, E., and Schedl, P. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. Journal of molecular biology *185*, 341-358.

Venken, K.J., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747-1751.

West, A.G., Huang, S., Gaszner, M., Litt, M.D., and Felsenfeld, G. (2004). Recruitment of histone modifications by USF proteins at a vertebrate barrier element. Molecular cell *16*, 453-463.

Wood, A.M., Van Bortle, K., Ramos, E., Takenaka, N., Rohrbaugh, M., Jones, B.C., Jones, K.C., and Corces, V.G. (2011). Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. Molecular cell *44*, 29-38.

Zhao, K., Hart, C.M., and Laemmli, U.K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. Cell *81*, 879-889.

Zhong, X.P., and Krangel, M.S. (1999). Enhancer-blocking activity within the DNase I hypersensitive site 2 to 6 region between the TCR alpha and Dad1 genes. J Immunol *163*, 295-300.

Zhu, X., Ling, J., Zhang, L., Pi, W., Wu, M., and Tuan, D. (2007). A facilitated tracking and transcription mechanism of long-range enhancer function. Nucleic acids research *35*, 5532-5544.

APPENDIX
TABLES OF PRIMERS AND PRIMTER SEQUENCES

Primer Name	Primer Sequences (5'⟶3')	Purposes
P3-(delM2/attRL)-5'	CTGAGTGAGACAGCGATATGA	Check if M2 or *attR*/*attL* are removed
yel-wing-enh-(delattRL)-3	AATGTGGTAAT GGGCGATAAT	Check if *attR*/*attL* is removed
LS4D11-5'	Phosphate-GATCTTTAGAGCACTATTCAATAATTC TCTTGATTTCAGTCACGATATTCTTCCACCAACC GATAGTATCGCACACG	Make double stranded LS4D11 oligo
LS4D11-3'	Phosphate-GATCCGTGTGCGATACTATCGGTTGGT GGAAGAATATCGTGACTGAAATCAAGAGAATT ATTGAATAGTGCTCTAAA	Make double stranded LS4D11 oligo
4D21-*Bgl*Spe-5'	gatctGTTTAGAGCACTATTCAATAATTCTCTTGAT GAACTA	Make double stranded LS4D21oligo
4D21-*Bgl*Spe-3'	ctagTAGTTCATCAAGAGAATTATTGAATAGTGCT CTAAACa	Make double stranded LS4D21oligo
LS4-D-min-*BamBgl*-5'	gatctGCACTATTCAATAATTCTCTTGTCACGATAT TCTTCCACCAACCGATAGTATCGg	Make double stranded LS4Dmin oligo
LS4-D-min-*BglBam*-3'	GATCCCGATACTATCGGTTGGTGGAAGAATATC GTGACAAGAGAATTATTGAATAGTGCA	Make double stranded LS4Dmin oligo
fp-scs'-M-5'	TCTGCTTCAGTAAGCCAGATGC	Amplify M fragment without label
fp-scs'-M-3'	AGTTAGCTCACTCATTAGGCAC	Amplify M fragment without label
fp-scs'-FAM-5'	FAM-TCTGCTTCAGTAAGCCAGATGC	Amplify fluorescent labeled M fragment
fp-scs'-HEX-3'	HEX-AGTTAGCTCACTCATTAGGCAC	Amplify fluorescent labeled M fragment
Scs'-M-5'del-5'-*Bgl*II	CCAGATCTAAGCTTTTAGAGCACTATTCAATAA TTCTCTTG	Delete 5' end of M fragment
Scs'-M-5'del-3'-*Bam*HI	GTGGATCCGTTCGTTTGAATTGTGAAGC	Delete 5' end of M fragment

(table continued)

Primer Name	Primer Sequences (5'⟶3')	Purposes
Scs'-M-3'del-5'-*Bgl*II	CCAGATCTAAGCTTTGTATTTGTATAGAGATAG AAATTAAGG	Delete 3' end of M fragment
Scs'-M-3'del-3'-*Bam*HI	GTGGATCCGTGTGCGATACTATCGGTTG	Delete 3' end of M fragment
Scs'-LS4-D-5'-*Bgl*II	CCAGATCTAAGCTTTTAGAGCACTATTCAATAA TTCTCTTGATTTCAGTCACGATATTC	Make M35spacerΔ fragment
Scs'-LS4-D-3'-*Bam*HI	GTGGATCCGTGTGCGATACTATCGGTTGGTGGA AGAATATCGTGACTGAAATCAAGAG	Make M35spacerΔ fragment
Scs'-LS4-spacer-D-5'-*Bgl*II	CCAGATCTAAGCTTTTAGAGCACTATTCAATAA TTCTCTTGATTTCCAACAGGTAAGACAGTTCGC AGGTAATAACTAG	Make M35spacer* fragment
Scs'-LS4-spacer-D-3'-*Bam*HI	GTGGATCCGTGTGCGATACTATCGGTTGGTGGA AGAATATCGTGACTATGCAGGCACTAGTTATTA CCTGCGAACTGTC	Make M35spacer* fragment
Syn-M-5'-27	GGCTCGTATGTTGTGTGGAATTGTGAG	Sequencing pSyn series plasmids
Syn-M-new-3'	CAGATGCTACACAATTAGGCTTGTAC	Sequencing pSyn series plasmids
LS4-52'	TaactagtgcCTgcatATTTCAAATTGAAATAATACAC A	Introduce LS4 mutation
LS4-32'	gcactagttattacctTCTAAACTTTGGCATTTATGTTA	IntroduceLS4 mutation
C4scs-5'	ATGTCCGTGGGGTTTGAATTAACT	Sequencing pC4-scs-*attB* series plasmids
C4scs-new-3'	ACATACATACTAGAATTCGGTACCCGC	Sequencing pC4-scs-*attB* series plasmids
attR-5b'	agcttCTCGAGGTAGTGCCCCAACTGGGGTAACCT TTGGGCTCCCCGGGCGCGTACTCCACcatg	Make double stranded *attR* sequence
attR-3b'	GTGGAGTACGCGCCCGGGGAGCCCAAAGGTTA CCCCAGTTGGGGCACTACCTCGAGA	Make double stranded *attR* sequence

(table continued)

Primer Name	Primer Sequences (5'⟶3')	Purposes
attL-NsiI-5'	tgcatCGGTGCGGGTGCCAGGGCGTGCCCTTGAGT TCTCTCAGTTGGGGGCGTAGgcatgcatgca	Make double stranded *attL* sequence
attL-*Sph*I-3'	tgcatgcCTACGCCCCCAACTGAGAGAACTCAAGG GCACGCCCTGGCACCCGCACCGatgcatgca	Make double stranded *attL* sequence
yel-*sph*-5'	tcacgcatgcGACTATTAAATGATTATCGCCCG	Clone *yellow* gene
yel-*sph*-3'	CACTGCATGCCTTTCCCTGCACCCAAAC	Clone *yellow* gene
TRW-*Sal*I-3'	GATCGTCGACAGACATGATAAGATACATTGATG AGTTTG	Amplify mRFP-fibrillarin and EGFP-Nop56
TRW-*Xba*I-5'	TCGAGGCCTGTCTAGAGAAG	Amplify mRFP-fibrillarin and EGFP-Nop56

VITA

Yu Ge was born in Henan, China. He is the son of Tianwen Ge and Xiaoping Wang. He obtained his high school degree from Huixian No. 1 High School, Henan, China in June 2005. He graduated with a Bachelor of Science degree in biological science from Sichuan Agricultural University, Sichuan, China, in June 2009. Mr. Ge started to pursue a Philosophy of Doctor degree in biological science in Louisiana State University, Baton Rouge, USA, in August 2009. Then he decided to do a Master of Science program rather than Philosophy of Doctor program in the same department. Mr. Ge is going to graduate with a Master of Science degree in biological science from Louisiana State University in August 2012.